PRESSES UNIVERSITAIRES DE VALENCIENNES

European Association of Cognitive Ergonomics

Seventh European Conference on Cognitive Science Approaches to Process Control

Septième Conférence Européenne sur les Approches Cognitives du Contrôle de Processus

PROCEEDINGS
ACTES

CSAPC'99

Villeneuve d'Ascq, F, 21-24 Sept. 1999

A EACE conference series (European Association of Cognitive Ergonomics)

Une série de conférences de EACE, Association Européenne d'Ergonomie Cognitive

LEZ VALENCIENNES
1999 N° 28

Mainly sponsored by :
Principalement soutenue par :

LAMIH
LABORATOIRE D'AUTOMATIQUE ET DE MECANIQUE INDUSTRIELLES ET HUMAINES

CNRS

GROUPE UVHC
Valuval

UVHC
UNIVERSITE DE VALENCIENNES ET DU HAINAUT - CAMBRESIS

CSAPC'99
COGNITIVE SCIENCE APPROACHES TO PROCESS CONTROL
APPROCHES COGNITIVES DU CONTRÔLE DE PROCESSUS

VILLENEUVE D'ASCQ, F, 21-24 SEPT. 1999

ORGANIZING COMMITTEE / COMITE D'ORGANISATION

J.M. HOC (CNRS-UVHC, F)
P. MILLOT (CNRS-UVHC, F)
E. HOLLNAGEL (OCDE, N)
P.C. CACCIABUE (JRC, I)

PROGRAM COMMITTEE / COMITE DE PROGRAMME

R. Amalberti (IMASSA, F)
S. Bagnara (U. Sienne, I)
R. Beuscart (CERIM, F)
G. Boy (EURISCO, F)
G. De Michelis (U. Milano, I)
M. Frese (U. Giessen, D)
D. Harris (Cranfield., UK)
P. Jones (U. Illinois, USA)
M. Lind (Techn. U., DK)
N. Moray (U. Surrey, UK)
L. Norros (VTT, SF)
J. Patrick (Cardiff U., UK)
B. Pavard (CNRS, F)
F. Saad (INRETS, F)
O. Stock (IRST, I)
O. Svenson (U. Stockholm, S)
M.C. Thomas (CNRS, F)
T. van der Schaaf (TU Eindhoven, NL)
G. van der Veer (U. Amsterdam, NL)
P. Wieringa (TU Delft, NL)
H. Yoshikawa (Kyoto U., J)

Foreword

This publication group together the texts of the presentations made at CSAPC'99, the Seventh European Conference on Cognitive Science Approaches to Process Control, which took place in Villeneuve d'Ascq (France) from 21st to 24th of September 1999.

CSAPC is one of the two conference series organised by the European Association of Cognitive Ergonomics (EACE), the other one being the European Conference on Cognitive Ergonomics (ECCE). The first CSAPC conference was held in 1987 near Paris and the subsequent conferences have taken place in Denmark, Finland, France, Italy, and United Kingdom. This biennial conference series has successfully brought together European experts in the fields of cognitive psychology and ergonomics, human-machine systems, and artificial intelligence, to discuss multidisciplinary research in the design and evaluation of complex, dynamic human-machine systems. This research has made important contributions to a number of application domains such as industrial process control (e.g., nuclear power plant or steel industry), aviation (e.g., glass-cockpit or air traffic control), ship navigation, car driving (e.g., advanced computer support), train traffic, and medicine (e.g., anaesthesiology in operation room). The CSAPC series of conferences supplement more the specialised meetings within each of these applications, by addressing the common research topics and problems at a level corresponding to solutions from cognitive ergonomics and engineering. This focus on similarity and differences across application domains is one of the main objectives of CSAPC.

The theme of CSAPC '99 was ***Human-Machine Reliability and Co-operation***. This theme was chosen because issues of reliability and co-operation are fundamental to understanding the risks that inevitably exist in complex human-machine systems. This can be exemplified by the question: *Under what conditions will the human be able to control the machine and, how can the overall human-machine system apply its adaptive capabilities to unexpected situations?* The presentations were organised in seven sessions, the first four dealing more especially with co-operation, the following two with human reliability, and the last one grouping together some presentations of general interest.

The presentations on co-operation successively approached the issue from the point of view of human-machine interfaces (support provided to the operator by information presented) and of computer assistance (support by an active machine in the task). The elements of a cognitive approach to co-operation resulted in defining some requirements to be satisfied by machines when one wants to qualify them as co-operative machines. Finally, some presentations dealt with human-human co-operation, which is more and more a typical feature of complex and dynamic situation management.

Human reliability has been accounted for by error analysis and management. The error analyses presented referred to the large corpus of incidents and accidents that exists, as well as to more detailed individual protocol analyses. Error management was tackled from very diverse points of view, at the prevention, recovery, or prediction level.

Finally, the general interest presentations addressed issues on temporal, cognitive, and social aspects of dynamic situations and methodological reflections of common interest.

Jean-Michel Hoc
Patrick Millot
Erik Hollnagel
Pietro Carlo Cacciabue

Cette publication rassemble les textes des communications présentées à CSAPC'99, la Septième Conférence européenne sur les Approches cognitives du Contrôle de Processus. Cette conférence s'est tenue à Villeneuve d'Ascq (France) du 21 au 24 Septembre 1999.

CSAPC est l'une des deux séries de conférences organisée par EACE (l'Association Européenne d'Ergonomie Cognitive), l'autre étant ECCE, la Conférence européenne d'Ergonomie cognitive. La première conférence CSAPC a été organisée en 1987 près de Paris et les suivantes se sont tenues au Danemark, en Finlande, en Italie et au Royaume-Uni. Cette conférence biennale vise à rassembler des experts européens reconnus dans les domaines de la psychologie et de l'ergonomie cognitive, des systèmes homme-machine et de l'intelligence artificielle, pour partager leurs résultats de recherche, dans une perspective pluridisciplinaire, pour la conception et l'évaluation des systèmes homme-machine complexes et dynamiques. Ces recherches ont un impact important dans un grand nombre de domaines d'application, aussi divers que le contrôle de processus industriel (ex. : centrales nucléaires ou sidérurgie), l'aéronautique (ex. : glass-cockpit ou contrôle de trafic aérien), la navigation maritime, la conduite automobile (ex. : aides informatiques avancées), le trafic ferroviaire et la médecine (ex. : anesthésie en salle d'opération). La série CSAPC prolonge les colloques spécialisés dans chacun de ces domaines d'application, en traitant des sujets et des problèmes de recherche communs à un niveau adapté pour trouver des solutions en ergonomie ou en génie cognitif. Cette concentration sur les ressemblances et les différences d'un domaine d'application à l'autre est l'un des principaux objectifs de CSAPC.

Le thème de CSAPC'99 était **La fiabilité et la coopération homme-machine.** *Ce thème a été choisi car les questions de fiabilité et de coopération sont fondamentales pour comprendre les risques qui sont inévitables dans les systèmes complexes. Il peut être illustré par la question : À quelles conditions l'opérateur humain sera-t-il en mesure de contrôler la machine et comment le système homme-machine, dans son ensemble, pourra-t-il appliquer ses capacités adaptatives à des situations imprévues ? Les communications ont été regroupées dans sept sessions, les quatre premières portant plus particulièrement sur la coopération, les deux suivantes sur la fiabilité humaine et la dernière regroupant quelques contributions d'ordre général.*

La coopération a été abordée successivement du point de vue des interfaces homme-machine (soutien apporté à l'opérateur par les informations présentées), puis des assistances informatiques (soutien grâce à une machine active dans la tâche). Les éléments d'une approche cognitive de la coopération ont permis de tracer quelques exigences auxquelles devraient répondre les machines pour être qualifiées de coopératives. Enfin, sur ce même thème de la coopération, quelques communications ont porté sur la coopération entre humains qui est de plus en plus un trait caractéristique de la gestion des situations complexes et dynamiques.

La fiabilité humaine a été abordée au travers de l'analyse de l'erreur et de la gestion de l'erreur. Les analyses d'erreurs rapportées se sont surtout appuyées sur de larges corpus d'incidents et d'accidents, mais également sur des approches plus fines avec des analyses de protocoles individuels. La gestion des erreurs a été abordée selon des points de vue très divers, au niveau de la prévention, de la récupération et de la prédiction des risques.

Quelques communications d'intérêt général ont conclu cette conférence sur les aspects temporels, cognitifs et culturels des situations dynamiques et sur des réflexions d'ordre méthodologique.

Jean-Michel Hoc
Patrick Millot
Erik Hollnagel
Pietro Carlo Cacciabue

Session 1: Computer interface
Les interfaces informatiques

Session 2: Computer assistance
Les assistances informatiques

Session 3: Cognitive cooperation
Coopération cognitive

Session 4: Human-human cooperation
Coopération homme-homme

Session 5: Error analysis
L'analyse de l'erreur

Session 6: Error management
La gestion de l'erreur

Session 7: Complex and dynamic systems
Les systèmes complexes et dynamiques

Session 1: Computer interface
Les interfaces informatiques

Information Types and Mapping in Process Displays

Chr. Rud Pedersen
Seven Technologies
Bregnerødvej 94
DK-3460 Birkerød, Denmark
e-mail: crp@sevent.dk

ABSTRACT

It is argued that explicit knowledge of information types and possible presentations of this information are needed. A categorisation, which maps information types to graphical modalities (visualisation types), is proposed. The reasons to create, and the benefits of, the categorisation are: 1) The categorisation will aid the display builder in selecting the relevant presentation for given information; it becomes obvious what the visual possibilities are. 2) The categorisation will work as a frame for analysing existing displays. Such an analysis will explicitly reveal the information content in the display, which is required in order to assess the display with regard to which operator tasks it is suitable for.

The proposed categorisation should be regarded as an attempt to raise an issue rather than to present the ultimate solution.

KEYWORDS

Information Mapping, Process Control, GUI, Categorisation, Information Types, Graphical Modalities.

INTRODUCTION

Today, industrial process displays are often designed in an ad hoc manner based on piping and instrumentation diagrams. Trend curves are the most common presentation used to visualise the development of process variables over time. It is very seldom that abstract concepts such as goals and process operations are shown in the process displays. The graphical elements are selected without giving much consideration to what information is mediated to the operator and what other possibilities exist for presenting this information.

Knowledge about information types and possible visualisation means is often implicit in the decisions made by the display builder. The focus in this paper is therefore to make this knowledge explicit by identifying and classifying the information types found in process displays. Combining the abstract classes of information with graphical modalities makes it possible to explicitly state the possibilities of visualising given information.

Bertin (1983) identified maps, networks, and diagrams as basic graphical types. Moreover, Bertin identified the following visual dimensions[1]: shape, position, size, orientation, colour (hue), brightness and texture. Bertin did not have a specific work domain in focus.

Tufte (1990) describes different ways of presenting information using examples from all kinds of working domains. The examples are grouped into the following headings: micro/macro readings, layering and separation, small multiples, colour and information, and narratives of space and time. Many inspiring examples are provided in Tufte but the design problem of how given information should be presented is not dealt with. However, guidelines for the use of colours, layers and separation are given through the examples. Hence the grouping of visualisation examples given by Tufte does not provide the methods needed to systematically design user interfaces in process control.

Neither Bertin nor Tufte deal with the working domain of process control. Their main focus is on how to visualise data variables (later on referred to as magnitudes) and relations between data variables, usually in graphs. None of them treat the real-time aspect involved in process control.

Basically, the visualisation problem can be separated into visualisation of entities, i.e. both (touchable) physical items and (imaginary) concepts, and into visualisation of the relations between entities. This is especially the case in process control where physical items such as plant components and concepts like process operations and alarm (states) are important information types.

---

[1] Bertin (1983) refers to visual dimensions as visual variables and uses the term value instead of brightness.

IDENTIFIED INFORMATION TYPES

The existing displays regarded in the analysis of information types are: mimic diagrams, trend curves, Paulsen's overview display (Paulsen, 1996), the mass data display (Beuthel et. al., 1995), the ecological interface (Vicente and Rasmussen, 1990) and Goodstein's functional display (Goodstein, 1985).

The result is shown in Table 1. The following is a description of the working procedure and clarification of the terminology and the chosen categorisation.

The method to identify the different information types in process displays has been a combination of bottom-up and top-down approaches. Bottom-up by analysing existing process displays, which gave information types such as component status, magnitudes (numerical value) of process variables, alarm states, set points, deviation from normal condition and connections between components. Top-down by leaving the specific displays and considering abstract categories of information. The information types derived from the analysed displays where then reorganised into these abstract information categories.

Deciding and defining the number of levels in such abstractions call for many considerations and discussions. The levels shown in this paper are based on the analysis of existing displays and the author's subjective assessment of what is appropriate and practical for such a categorisation of information types. Hence this should be regarded as a first attempt to construct such a categorisation. Therefore, the number of abstractions among the content is open for discussion, though it is believed that the proposed categorisation encompasses a great part of the information types needed in process displays and that the chosen levels of abstraction are suitable for structuring this information.

The following generic information types were identified and defined:

Process variables, i.e. continuous sensory signals, e.g. 4 – 20 mA.

Status indicators, i.e. discrete component signals, e.g. opened or closed for a valve.

Intervention points, i.e. the physical devices used by the operators to execute process operations. Examples are command buttons for starting process operations, input boxes for entering parameter values or sliders or spin buttons for adjusting set points.

- Magnitudes (numeric values)
  - Process variables
    - Actual value
    - Time series

    - Absolute value
    - Relative value
      - to admissible range
      - to typical range
      - to normal value

    - Division into classes/intervals
  - Status indicators
    - Actual value
    - Time series

    - Deviation from normal
- Physical Items
  - Components
    - Properties
  - Intervention points
    - Activation of process operation
    - Activation of component
    - Parameter adjustment
- Concepts
  - States (including alarms)
    - Actual value
    - Time series

    - Deviation from normal
  - Process Operations
    - Operator controlled
    - Automated
  - Functions
    - Design constraints
    - Constraints from laws of physics
  - Plant goals
    - Set point

    - Production
    - Safety
- Relations
  - Process variable – process variable

  - Status indicator – status indicator
  - Status indicator – process variable

  - State – process variable
  - State – status indicator
  - State – process operation
  - State – plant goal

  - Component – process variable
  - Component – status indicator
  - Component – component
  - Component – function
  - Component – process operation
  - Component – plant goal

  - Process operation – process variable
  - Process operation – status indicator
  - Process operation – function
  - Process operation – process operation
  - Process operation – plant goal

  - Plant goal – plant goal

  - Intervention point – process variable
  - Intervention point – status indicator
  - Intervention point – component
  - Intervention point – process operation
- Events
  - Progress
  - Sequence
- Location
- Language

Table 1. Identified information types placed in categories.

States, i.e. a derived signal based on several components and the state of the control system. E.g. the state of the process operation: filling can be starting, running, stopping and interlocked. A temperature sensor can be in different alarm states. Other types of states such as e.g. a tank level divided into: full, normal, nearly empty and empty also exist.

Components, i.e. the actual physical devices in the plant, both passive and active, e.g. pumps and temperature sensors (active components) or pipes and storage tanks (passive components).

Process Operations, i.e. the operations that the operators can perform on the plant to achieve the plant goals.

Functions, i.e. the design intentions made for the plant. E.g. cooling to x °C or transportation of y hl/h (see Pedersen and Lind, 1999, for definition and distinction between process operations and functions).

Plant Goals, i.e. the goals related to production, quality, safety, etc. when then plant is in operation.

Relations, i.e. possible correlation or connection between entities.

Both status indicators and states are introduced to distinguish between information based on (discrete) signals available directly from the process plant (status indicators) and information derived and assessed from the plant signals (states).

Based on the information found in the existing displays the generic information types were specified in further detail, as shown in the subclasses in Table 1. Moreover, the generic information types were abstracted into superior categories. Both the specification and abstraction will be commented in the following.

For *process variables, status and state indicators* either the actual value or time series can be shown. The time series can be either the historic values or predicted values; though predicted values were not found in any of the analysed displays. For process variables, the absolute or relative value can be shown. A value can be relative to the admissible range, the typical range or the normal value (deviation from normal). These absolute or relative values can be shown either as actual values or time series. Finally, process variables can be divided into classes, that is process variable states. For example, a pump's speed is divided into high, medium and low or the well-known alarm states for sensory input: high alarm, high warning, normal, low warning and low alarm. The abstract class encompassing process variables and status indicators is magnitudes (numeric values).

Components and intervention points are placed in the abstract category of physical items as opposed to concepts. A suggested subtype to components is properties, though not found in any of the analysed displays. It is placed in the categorisation because properties, e.g. manufacture, might be useful for maintenance tasks. Intervention points are divided into activation of process operations (control through an automated system), activation of components (each component is individually controlled by the operator) and parameter adjustment, e.g. a set point or selection of a tank number to be filled.

The *non-physical related information that is states, process operations, functions, and plant goals* are placed in the broad category of concepts. Process operations are divided into two. Operator-controlled and automated process operations. A clear distinction is made between intervention points to process operations and the mediation of the process operation themselves, i.e. which process operations exist and what are their relations and states. Functions describe the constraints exhibited by the design solutions or the constraints made by the laws of physics. Plant goals can be related to the production including economy and quality or the safety aspect of the plant. Another type of goal is set points for controllers. Again notice the distinction between a set point as a goal (the temperature should be 80 °C) and the means to adjust the set point as an intervention point.

For states, process operations, functions and goals, the entire plant can be regarded as a whole or decomposed into subsystems. The decomposition is needed to describe a plant in detail and does not influence the identified information types or how they should be visualised. However, the decomposition provokes a problem of navigation between different levels of plant descriptions. The navigational problem is outside the scope of this paper.

Relations as information types are important but difficult to categorise. They are important because it is the relations between artefacts that actually carry most of the information needed to act in a process plant or, more philosophically, in the world. A pump is not useful if it is not connected to pipes and tanks and it is not possible to operate it without knowing the relation between the start button and the pump. This also illustrates the difficulties with relations because many relations are so implicit and well-known that they are not regarded as relations (of course the button starts the pump). In theory, an innumerable amount of relations exist and to make a useful categorisation of information types, a limited number of relations must be selected. Table 1 shows the relations found in the analysed displays together with others, relevant in the context of controlling a process plant.

Events are another type of information, which deal with the dynamics of the process plant. Two subtypes are proposed: progress and sequence. Progress is related to the time series of process variables, status and state indicators, whereas sequence relates to execution of process operations either by the operator or the automation system.

Location is a fundamental information type related to the physical placement of objects. It can be argued that a location is a property of a physical item (e.g. a plant component). Here location is placed in a separate class because a specific graphical modality is used to visualise a location, which is different from the graphical modality used to visualise a physical item.

Language is the last fundamental information type and it is connected to procedural information and messages.

GRAPHICAL MODALITIES

Some modalities are identified by May independent of the media (see Pedersen and May, 1998). In this paper, regarding process displays, a limitation to the visual media is made. The modalities are the same across all media, but the medium becomes important when a modality's ability to present information is assessed, in the next section.

The modalities from Pedersen and May, 1998, are listed in Table 2 and a description of the invariant properties of the modalities is added. The invariant properties are the features that make the modality unique. Notice that almost all useful graphics are combinations of several modalities.

INFORMATION TYPES AND GRAPHICAL MODALITIES

From the invariant properties of the modalities in Table 2, it is obvious that some modalities are more suitable to present given information than others. For example, a map is very good at presenting a location whereas a symbol cannot be used to inform about a location.

In Table 3, the abstract categories of information types are mapped and assessed to the graphical modalities.

Through the analysis it is noticed that the image, symbol and text modalities are suitable to present items and concepts as opposed to maps, graphs, structural and conceptual diagrams, which mediate relations.

Modality	Invariant properties
Image	Resemblance to a physical item.
Map	Size, location, shape, and orientation maintained relative to a background, often earth.
Graph	Magnitudes mapped into space. E.g. Trend curves or bar graphs.
Structural Diagram	Size, location, shape, and orientation maintained. E.g. engineer's construction drawings.
Conceptual Diagram	Relations between items, instead of size, location and orientation. E.g. flow charts or networks.
Symbol	Simplification and recognition of entities by conventions. Examples of entities are physical items, concepts, and events.
Text	Alphabetic characters (symbols)

Table 2. Graphical modalities and invariant properties.

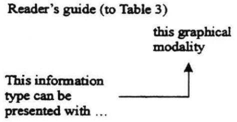


	Image	Map	Graph	Structural Diagram	Conceptual Diagram	Symbol	Text
Magnitude			●			**O**	●
Physical Items	●					●	O
Concepts						●	**O**
Relations		●	**O**	●	●		O
Events			**O**		**O**		O
Location	O	●		**O**			O
Language						O	●

Legend: ● =good, **O**=acceptable, O=bad, = not possible

Table 3. Analysis and assessment of mappings between identified abstract information types and graphical modalities.

Text is the only graphical modality able to communicate all the information types, though it is not always the best choice (e.g. relations between several plant components are not easily described with text only). Further, text and image are the only graphical modalities that can be used individually. The other graphical modalities need to be combined with other modalities, often as a minimum with text.

All the information except events is communicated in a good (reasonable) way by one or more of the identified graphical modalities. The problem with events is that time is involved. The dimension of time can either be mapped into space using the graph modality or it can be animated. Sequences in time can be visualised by use of conceptual diagrams or described with text.

EXAMPLE

Suppose that the relations between physical items (components) and concepts (process operations) should be mediated. Maps, structural and conceptual diagrams are the graphical modalities that are good at expressing relations. A map or structural diagram is not appropriate here because a concept is involved and concepts do not have a location (a map or structural diagram would be relevant if relations between physical items should be mediated). Therefore, a conceptual diagram is chosen.

The physical items and the concepts themselves must be visualised. Symbols are chosen for the physical items and texts are chosen for the concepts.

The invariant or special feature of conceptual diagrams is that all the visual dimensions (c.f. Bertin, 1983) can be used to express the relations. The most obvious way to visualise the relation is by a connection line or by colour coding.

The example of a storage tank (see Pedersen and Lind, 1999) is suitable for illustrating some of the possible ways to visualise the components, the process operations and the relations between these physical items and concepts.

The term *dynamic relation* is used to emphasise the dynamic behaviour of the plant in operation as opposed to the *static relation* between the components. Of course, the plant can be reconfigured and the relations between the components change, but this is not of interest for the scope of operating the plant.

In Figure 1 colours (here highlighted with bold instead of colours) are used to mediate the dynamic relations between components and process operations. For each process operation, the components involved are shown on an individual mimic diagram (which is a conceptual diagram embedded with symbols).

In Figure 2, connecting lines are used to mediate the dynamic relations only, meaning that the static relation in form of the mimic diagram is omitted.

In a real plant such a view becomes very confusing because of the one-to-many relations and because of the number of components and process operations. The one-to-many relations is the main problem, which also makes colour coding difficult. How should several colours be applied to one graphical item? An attempt to use transparent colours is shown in Figure 3 on the next page.

The one-to-many relation also makes it impossible to use the position to mediate relationships using the gestalt principles of proximity. One graphical item cannot be placed near several others without copying it.

The static relation between the components, shown by the piping and instrumentation diagram, is very useful as a background for the dynamic relations such as e.g. relations between components and process operations. It is also seen that the display easily becomes very crowded when more relations are added. This can be handled by letting the operator or the automation system control when the dynamic relations shall be shown.

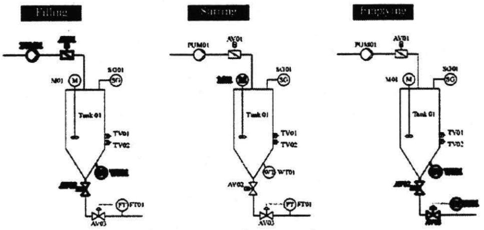


Figure 1. The dynamic relations between process operations and components are shown on the mimic diagrams.

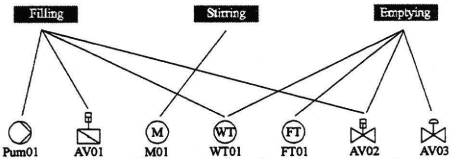


Figure 2. Mediating only the dynamic relations between components and process operations by use of connecting lines.

Figure.3. Showing one-to-many relations by use of colour coding (here grey scales).

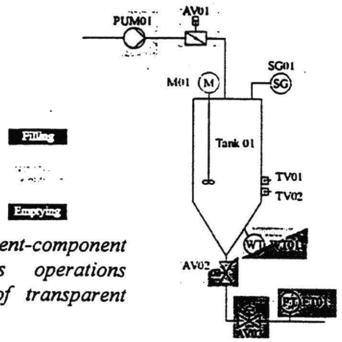


Figure.4. Mediating static component-component and dynamic component-process operations relations simultaneously by use of transparent colour coding (here grey scales).

Finally, the colour coding of the one-to-many relations is used in the mimic diagram making it possible to view both dynamic and static relations simultaneously. To make this simple example a bit more realistic, it is assumed that the bottom valve AV02 must be closed during stirring. This is to get an idea of how the transparent colour coding might work when three process operations use the same components. The result is shown in Figure 4.

From Figure 4 it is seen that it is possible to code at least 3 entities by use of transparent colours. The static relations are still visible in the background though the focus is on the coloured dynamic relations. Whether this is useful in a real plant, where the density of components and process operations is higher, or whether it will appear as a mishmash of colours remain to be investigated.

CONCLUSION

Information types are identified in process displays. The information types are organised in the following main categories: magnitudes, physical items, concepts, relations, events, locations and language.

The main categories of information types are mapped and assessed to graphical modalities. From this mapping it is concluded that the image, symbol and text modalities are suitable to present entities as opposed to maps, graphs, structural and conceptual diagrams which mediate relations between entities.

From the examples of visualisation techniques developed and discussed, it can be concluded that the mappings between information types and graphical modalities (Table 3) are useful as a tool forcing the display designer to make the design choices explicit and as a tool for structuring the analysis of existing process displays.

REFERENCES

Bertin, J. (1983). *Semiology of Graphics*. University of Wisconsin Press, Madison

Beuthel, C., Boussoffara, B., Elzer, P., Zinser, K. and Tissen, A.. (1995). Advantages of Mass-Data-Display in Process S&C. *6th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems*. Massachusetts Institute of Technology Cambridge, MA, USA. June 27-29, 1995

Goodstein, L. P. (1985). *Functional Alarming and Information Retrieval*. Risø National Laboratory. Risø-M-2511

Paulsen, J. L. (1996). Dynamic Overview Displays for Process Plants. *Proceedings Cognitive Systems Engineering in Process Control CSEPC 96*, Kyoto Japan

Pedersen, C. R. and Lind, M (1999). Conceptual Design of Industrial Process Displays. To be published in *Special Issue of Ergonomics, Time and Space in Process Control.*

Pedersen, C. R. and May, M (1998). Visualisation in Process Control. *7th IFAC Symposium on Analysis, Design and Evaluation of Man-Machine Systems*. Kyoto, Japan. September 16-18 1998.

Vicente, K. J., Rasmussen, J. (1990). The ecology of Human-Machine Systems II: Mediating "Direct Perception" in Complex Work Domains. *Ecological psychology*. 2(3), pp. 207-249.

Tufte, E. R. (1990). *Envisioning Information*. Graphics Press, Fourth printing, July 1994

RÉSUMÉ

TYPES D'INFORMATIONS ET MISE EN RELATION SUR LES ÉCRANS DE CONTRÔLE DE PROCESSUS

On présente une catégorisation qui met en relation des types d'informations à des modalités graphiques. Cette catégorisation est utile pour aider le concepteur et pour analyser des interfaces existantes.

MOTS CLÉS : Mise en relation d'informations, Contrôle de processus, GUI, Catégorisation, Types d'information, Modalités graphiques

SITUATION AWARENESS AND PROCEDURE FOLLOWING

Gabrielle de Brito and Guy Boy
debrito@onecert.fr

EURISCO
European Institute of
Cognitive Sciences and Engineering
4 avenue Edouard Belin, 31400 Toulouse
+33 (0) 5 62 17 38 38

Université René Descartes
Laboratoire d'ergonomie
informatique de Paris V
Rue des Saints-Pères,
75006 Paris

ABSTRACT

Aeronautics procedures are used as prescribed action lists to help human operators remember and follow mandatory steps that guarantee safety, workload and performance criteria. A study of the use of procedures in the civil aviation domain surveyed 207 pilots using four investigation methods, including the observation of 140 hours of full-flight simulator. The results of this study are used to address why human operators of safety-critical systems use, misuse or do not use procedures to keep control of a situation, and how they cope with situation awareness. This paper suggests that new perspectives on design may be required to support the further development of warning systems, the design of procedure and the definition of the pilots' role.

Keywords

Empirical findings, procedure following, situation awareness, safety-critical systems.

INTRODUCTION

Situation awareness of the human operator is a major issue in the control of current safety-critical systems, such as aerospace systems. It is also a fact that the control of most of these systems is commonly constrained by the use of procedures that are justified by at least four reasons: (i) to enable human operators to face increasing system complexity; (ii) to improve the coordination among human and machine agents; (iii) to support training; (iv) to provide a legal referent.

If the necessity of checklists seems today to be globally recognized by the civil aviation community, they do not agree about their content and their use (Gross, 1995) There seems to be a conflict between the expected use of checklists by manufacturers and airline companies, and the actual use of checklists by pilots (Degani & Wiener, 1990) The problem is that serious accidents involving human factors issues are caused most frequently by incorrect procedure following.

Accordingly, the French Civil Aviation Authority (DGAC) has financed a study at EURISCO concerning the use of onboard Airbus operational documentation in cockpits of new generation aircraft. In the United States, Degani & Wiener (1990; 1994) have focussed their study on the manner in which pilots use checklists in normal situations. We have completed this research by studying the use of the procedures in abnormal and emergency situations (systematically quantified analysis) while carrying out an analysis of pilots' cognitive activity.

The goal of this study[1] is to understand how pilots follow procedures dictated by management philosophy and policy embodied within flight deck procedures. The role of human operators induced by the use of procedures is not uniformly acknowledged in the aeronautics industry for example. Some manufacturers require that human operators consider procedures as prescriptions. Others require that they consider procedures as work support tools. This paper emphasizes a common cognitive contradiction between situation awareness and the role that is assigned to the human operator as a procedure follower. In particular, it examines the real role of the human operator. Indeed, there are two major hypotheses that are commonly put forward:

- H1: If the human operator executes tasks in a procedural mode and if procedures are well-designed, then the human operator does not need to be aware of the work situation (he/she is a procedure follower).
- H2: If the human operator needs to be aware of the work situation and if he or she needs to understand the procedures, then the human operator is a problem-solver.

---

[1] See Karsenty, Bigot & de Brito, 1995 ; de Brito, Pinet & Boy, 1998, 1999.

An experimental study was carried out to show the interrelated roles of human operators and procedures. The methodology that was used was targeted toward proving that pilots do not follow procedures and deducing why. This paper tries to extend the results of this study and proposes a framework of attributes that supports the reasons why pilots cannot be always considered as procedure followers, and the need to keep a well-maintained situation awareness. Recommendations for the design of more appropriate work aids will be provided.

WRITTEN OPERATING PROCEDURES USED IN AERONAUTICS

In the cockpit pilots have a large number of documents at their disposal. Onboard operational documentation consists of technical manuals on aircraft performance, flight maps, flight procedures to be used. All pilots have been extensively trained and have acquired specific knowledge for their use. Operational procedures can cover normal, abnormal and emergency situations. In abnormal situations, pilots encounter the procedure (more or less) for the first time. However, procedures for normal situations have a certain routine nature. To refer to Degani and Wiener (1990) "The principal difference ... lies in frequency of use. The abnormal checklist is very rarely performed by flight-crews during revenue flight; pilots are aware of its criticality, and very much aware that misuse or non-use of the abnormal checklist can transform a routine abnormality into an accident. The same cannot always be said about the normal checklist.".

Conditions for using operational documentation

Normal situations

Normal procedures define the basic flight scenario (divided into flight phases), assuming that all systems are functioning and are used correctly. The transition between flight phases is sanctioned by reading the paper "checklist", intended to control that vital preparation actions have been performed correctly. These checklists are on the back-cover of the Quick Reference Handbook (QRH) published by Airbus Industrie. The actions listed in these checklists are limited to respecting flight safety and efficiency and thus do not take into account the many basic control actions. For example, the checklist used during landing consists of three items whereas the associated normal procedure consists of 20 to 25 actions.

Abnormal and Emergency Situations

Abnormal and Emergency situations appear to be much more complex to deal with than normal situations, not the least being because they can be very different in nature. It would be almost worth processing each failure individually. Abnormal and emergency procedures are used in different and varying situations. Their number is enormous and the difficulty in *prescribing instructions* is due to the multiplicity of the situations. To reduce the occurrence of human errors, modern aircraft are designed to require as few actions as possible from the pilot in response to any failure or combination of them. These actions are prescribed in the normal and emergency procedures (called "dolists") which are designed to be followed precisely in order to recover the situation as quickly as possible. These dolists appear on a screen called the "Electronic Centralized Aircraft Monitoring" (ECAM) which presents 262 dolists for abnormal situations and 26 for emergency situations. Its main advantage is to allow direct interaction between the pilot and the system. Dolists appear automatically and indicate to the crew the nature of the failure. Moreover, as the actions are carried out, the lines containing the pertinent information disappear automatically, indicating that the action has been performed correctly. When the system can't detect the failure, pilots have at their disposal the Flight Crew Operating Manual (FCOM) which gathers all the procedures (normal, abnormal and emergency).

METHODOLOGIES

The theory behind this study is that procedure deviations can be explained by the pilots cognitive activity and the cognitive needs that this gives rise to. We will illustrate our remarks based on the results from the SFACT study using: (i) a analysis of the task and pre-analysis of possible deviations (ii) a questionnaire sent to ten airline companies of different nationality, (iii) observations in a full flight simulator with pilots of new generation aircraft (Karsenty, Bigot, & de Brito, 1995; de Brito, Pinet, & Boy, 1998; and de Brito, 1998) (iv) Group Elicitation Method sessions (Boy & Wilson, 1996)

Analysis of the task and pre-analysis of possible deviations. We started by analyzing the totality of procedures linked to the use of the checklists, and identified in parallel the totality of possible deviations linked to these procedures. The approach that we took to determine these deviations was inspired by the work of Hollangell, (1993) and the work of Reason (1990), aimed at defining a classification of erroneous actions according to their observable manifestation (or phenotype). This analysis served as a support for elaborating a questionnaire and an observation grid.

The questionnaire. The questionnaire contains 35 questions. The majority of questions consisted of (i) questioning pilots on certain deviations, (ii) justifying replies given. Where possible, closed questions or multiple choice were used to facilitate the reply. Ten companies using new generation Airbus were consulted and we received a return of 207 questionnaires out of the 606 sent. This relatively high figure reflects pilots' interest in the subject.

Simulator Observations. To compensate for the difficult access to real situations, all our observations were undertaken in mobile flight simulators (Full Flight Simulators) at the training center "Airbus Training" in Toulouse, with pilots in recurrent training. The possibility of observing pilots in the simulator presents the advantage of analyzing at close hand incident/accident situations.

Group Elicitation Method. Three Group Elicitation Method (GEM) sessions were used to capture consensus and contradictions among pilots on the use of checklists and procedures. During these sessions, pilots gave their viewpoints on the following question: "In the perspective of future electronic operational documentation, please give information concerning: (i) requirements (several layers, color, presentation, logic, details, mental load, etc.) (ii) problems and issues involved; (iii) right balance between paper and electronic documentation ».

The consecutive use of several methodologies seemed to us relevant in understanding the characteristics of following written operating procedures, and thus allowing an explanatory approach of the activity.

ANALYSIS AND DISCUSSION

The analysis, carried out using complementary methodologies, reveals that the large majority of pilots recognize the importance of operating procedures. Eighty percent of pilots questioned consider them to be important or very important. However, if we look at pilot satisfaction with current written operating procedures, the results are more mixed. Only 38 percent of pilots are satisfied with their contents, comprehensibility, presentation, size, adaptation to operation environment, ease of access to dolists. These results show that while pilots need procedures currently and they are not well adapted to their needs. Although cooperative crew work has positive effects, it only partially corrects for pilot error. This is why it is important to list the reasons for dissatisfaction with written operating procedures and to explain the most frequent deviations.

Pilots state that they generally use consistent behaviors when using written instructions, but at the same time admit to deviations from instructions generally in exceptional circumstances. They justify this behavior by three major classes: (i) the need to manage a different operational situation from what they envisaged (ii) loss of control of the activity, and (iii) the need to better understand the situation.

Manage a different operational situation

Flight deck procedures are pre-defined with respect to a class of situations. Each situation class can be very specific, sometimes unique, to general. As already seen, dolists used in abnormal situations may be executed when pilots judge that it is the right time. In contrast, emergency dolists must be started immediately. We have found that 92 percent of surveyed pilots think that aircraft safety is more important than initiating any dolist independently from the urgency of the situation. This point is clearly made to support dolist execution delays: only 39 percent of surveyed pilots claim that they never delayed a dolist. When a pilot (usually the captain) makes the decision to assure immediate safety of the aircraft, he or she is in change of the trajectory tracking task, i.e. speeds, heading, safety altitude, attitude, roll, yaw and navigation. Some pilots strongly advocate the fact that it would be crazy to risk an accident by being obliged to follow a written operating procedure when the normal trajectory of the aircraft is not appropriate or is degrading. Delaying the execution of a written operating procedure cannot qualify as a deviation but it comes as a result of a conflict resolution that pilots obviously needed to process. We then admit that pilots incrementally re-define emergency situations almost systematically. Consequently, intrinsic properties of real flight situations may make the expected applications of dolists partially or totally impossible. This is confirmed by the fact that pilots seems to feel a contradiction between the natural reactions to the emergency of the situation and procedure following that is necessarily slower. This contradiction may lead to the non-execution of the written operating procedure (at least at first), and instead of executing some items from memory. Memory-based execution can also be a reaction to the fact that information provided in the written operating procedure is inappropriate or incomplete. In addition to pilots' knowledge and know how, the central issue is the consistency between paper procedures and electronic procedures.

Paper procedures make some information explicit. In contrast, electronic procedures only present a list of actions to execute. It is the lack of information and the distance between the expected situation and the real situation that forces pilots to take action and incrementally construct in real-time the procedure that they need to execute.

Activity control loss

Normal checklists represent only a very small part of the rules and procedures that crews need to follow during a flight. Routine use, interruptions, and the operational context of the flight (radio, cabin crew, etc.) are often the main cause of activity control loss. In turn, activity control loss may be the cause of forgetting and poor cross-checking. Curiously, the aviation community tends to consider that such operational context disturbs checklist execution, whereas checklists are artifacts that actually disturb pilots' "natural" operations. It would be difficult to understand the relations between checklists and operations management without taking seriously into account time management. Time management is a constraint that comes from two types of requirements: commercial and technical. The commercial requirement constrains the pilot to satisfy a time schedule in a highly competitive environment, where air space is becoming overcrowded. The technical requirement constrains the pilots to a continuous control loop and a finite flight time where there is no room for waiting or thinking at leisure. Most decisions need to be made quickly. This is why experience, expertise and self-confidence are key factors for pilots.

The time constraint is crucial and pilots are permanently busy managing it. In addition, they need to manage another constraint dealing with procedure management. Procedures are designed to be executed linearly without interruptions. They are supposed to be an integrated part of a flight organized into flight phases of a standard sequential scenario, and are taken into account in the procedures.

Despite these major operational interference's leading to the non-application of prescribed procedures, 98 percent of pilots consider that checklists and dolists provide valuable during current flight management situation awareness assistance.

The resulting issue seems to be the scheduling of prescribed tasks. Problem solving under time constraints is a cognitive problem. For example, when an operational event forces the pilots to interrupt a checklist sequence, they use various strategies that depend on the duration of the interruption, or even its a-priori estimation. We found that forgetting items was more specifically due to the nature of the interruption. For example, when the ATC interrupts the pilots during the reading of a checklist, 68 percent of the pilots forget to complete the checklist appropriately. When the interruption is caused by a failure, 89 percent of pilots forget to complete the checklist.

Understanding the situation

As the pilot is ultimately responsible for failure management regardless of the nature of the warning system, it is important that he or she understands what the automation is doing, why it is doing this and what it intends to do next. To satisfy these commonly asked automation questions (Weiner & Curry, 1980) the design engineer must produce an interface concept capable of communicating this information effectively to the pilot. The same principle applies if the pilot is following instructions "'issued" by automation (as we find in the form of electronic checklists) where it is important that pilots understand what they are actually doing to the system (Hicks & de Brito, 1998). However as the failure sensing and diagnosis functions are already automated in modern aircraft, these variables are not normally known by the pilot, who without knowledge of the cause of the failure context may miss the effect of automated actions or fail to understand what they observe.

The lack of understanding of the actual situation and prescribed action rationale may lead pilots to use procedures in appropriately. It is difficult to infer from our results how many pilots execute prescribed actions without questions. For most pilots, understanding a prescribed action is constructing informational context that makes the situation relevant. This context may include (1) a sufficiently precise diagnosis that will highlight the interest of prescribed action when it is related to the desired state of the systems such as represented by pilots; (2) a more global representation of interconnections among systems capacities and situation awareness.

To summarize pilots need to understand what they need to do by using three cognitive functions according to their interest. These cognitive functions are used in real flight to compensate procedure limitations due to their design in a different context. More fundamentally however they are mandatory assets of pilots who need to keep control of the situation in all circumstances (Hollangell, 1993) Pilots need to know what they are doing all times and anticipate the consequences of their actions. It follows that many deviations from prescribed procedures are caused by a conflict between what is required and what the pilot needs to do to keep control of the situation. This may lead to serious consequences. We make the assumption that helping pilots maintain a satisfactory level of situational awareness and prescribed actions rationale, will contribute to reducing deviations that lead to serious consequences. The objective is thus to determine what, when and how pilots need to understand in order to act.

What do pilots need to understand?

It is however misleading to think that pilots should be aware of everything in order to use prescribed rules. System complexity and the amount of information that needs to be assimilated are very high and require prior definition of the level of understanding. Two possible approaches are: (i) determining what pilots need to understand before applying an abnormal or emergency procedure, (ii) determining what pilots are required to understand. This distinction is necessary since the pilot may not know the he/she needs certain information in order to fully understand the logic of a procedure, for example.

When and how do pilots need to understand?

Since flight situations are highly dynamic, it is almost impossible to categorize the level of situation awareness for all prescribed actions. Most actions can be started with a minimum level of situation awareness (Hunt & Rouse, 1981); situational awareness improves during actions and even after. It is thus necessary to provide the pilots with appropriate means that enable them to understand what is going on and what they should do. Pilots must be provided the means to understand these three different phases, which means dealing with each phase separately.

Understanding before the action. Training plays a vital role of providing pilots with the knowledge necessary to understand rapidly what is happening during a failure and what implications failure will have. However, as the current trend is to reduce the amount of initial training, this recommendation could be applied during recurrent training. Human-machine interfaces are also of importance in providing pilots with an up-to-date representation of system status (i.e., situation awareness) no matter what the situation.

Understanding during the action. This mode of understanding relies on the feedback from actions, but not just not simply the confirmation that an action has been carried out. Pilots wait for this feedback in order to better understand what is happening and to obtain a more global vision of the aircraft's status. Thus, compared with current practice, they require more informative feedback on possible consequences of an action on different levels.

Understanding after the action. This mode of understanding has the advantage of providing pilots with more time, which could allow getting additional information.

Given the possibility of using these three different phases, an essential question is to determine what information should be rapidly finished and what information can remain momentarily unknown without leading to a deterioration in pilots' performance.

CONCLUSION

The above results show that three main activities support and improve the situation awareness in safety-critical environments such as aerospace. These activities are recognized: situation management, control and understanding the current situation.

We have observed that onboard operation documentation provides a means of training and assisting users in carrying out their tasks. It must be efficient and easy to use. However, the contents of this type of documentation should be used to redesign human-machine interfaces better adapted to users. There is an important relationship between the interface and its procedures. The more an interface allows for easy system use, the less the user should require procedures; in a sense the interfaces affordance renders procedures unnecessary. Conversely, the more an interface is difficult to use, the more the user will require procedures. Thus, there is an industry trend toward specify procedures that are easier to use. Moreover, users often need to understand the rationale of suggested or required actions.

Through their training and experience pilots develop a mental model of aircraft systems. When pilots need to manage failures in aircraft systems, it is their mental models that guide their understanding of causality and expectations of effect. Encouraging the development of a mental model in a pilot that will give him/her a high degree of reasoning power in failure situations must be a primary objective of failure management training and for him/her to achieve this he/she must understand system automation behaviour (Hicks & de Brito, 1998). Designing a failure management system that is consistent with this principle and that complements this training must be an equally important objective.

It is important to recognize a link between a pilot's role in failure management, his/her task definition and his/her subsequent awareness of system status. Our study has shown that pilots do not behave currently according to the model of an "ideal" executant in abnormal and emergency situations: they seek to understand before acting. Beyond this general statement, a major conclusion of our work is that the majority of pilots interviewed acknowledge that understanding a prescribed action leads to representing the context that makes the action relevant. This context may include (i) the distance between the perceived state and the desired state of the systems that enables the pilot to give meaning to prescribed actions; (ii) a global representation of the interconnections among systems and flight objectives in order to evaluate possible consequences of a failure on the rest of the flight; (iii) other possible actions that, jointly exploited with the pilot's knowledge of system functioning and the operational environment status, will enable the pilot to judge the degree of relevance of the prescribed action.

In the future, the design engineer may limit the goal of interface technology and crew training to support a 'stripped down' rather than veridical mental model of automation, concentrating resources instead on designing a means of communicating an explanation of the automation's current actions, rationale and future intentions to the pilot. Indeed, the methodology of "cognitive function analysis" proposed by Boy (1998) offers a framework to enable a design team to anticipate and document cognition induced by both the design process and designed artifacts. It provides members of the design team with a common frame of reference and enables the team to share their views on the artifact being designed.

ACKNOWLEDGEMENTS

The authors wish to thank David Novick for his comments on this paper and Helen Wilson for her pecious help. The autors like to thank the pilots and airline companies who answered the questionnaire with such care and who accepted my presence in the simulator. The mass of information gathered is extremely valuable as a result.

REFERENCES

Boy, G., & Wilson, H. (1996). *Perspective on paper and electronic on-board documentation: Results of two GEM experiments with pilots* (Report No. T-96-041-V1). Toulouse, France: EURISCO Contractor Report.

Boy, G. A. (1998). *Cognitive Function Analysis*. London, England: Ablex Publishing Corporation.

de Brito, G. (1998). Study of the use of Airbus Flight deck Procedures and Perspectives for Operational Documentation. In G. Boy, C. Graeber, & J. M. Robert (Eds.), *International Conference on Human Computer Interaction in Aeronautics (HCI-Aero'98)* (pp. 195-201), Montreal, Canada: Ecole Polytechnique de Montréal: 27-29 May 1998.

de Brito, G., Pinet, J., & Boy, G. (1998). *Etude SFACT sur l'utilisation de la documentation opérationnelle dans les cockpits de nouvelle génération. Etude de l'interface Homme-Machine* (Report No. T-98-054). Toulouse, France: EURISCO Contactor Report.

Degani, A., & Wiener, E. (1990). *Human factors of flight-deck checklists : The normal checklist*: NASA contractor report 177549, Moffett Field, CA: NASA Ames Reasearch Center.

Degani, A., & Wiener, E. (1994). *On the design of flight-Deck Procedures*: NASA-Ames Research Center 177642. Moffett Field, CA: NASA Ames Reasearch Center.

Gross, R. (1995). *Studies suggest methods for optimizing checklist design and crew performance* : Flight safety foundation..

Hicks, M., & deBrito, G. (1998). Civil Aircraft Warning Systems: Who's Calling the Shots? In G. Boy, C. Graeber, & J. M. Robert (Eds.), *International Conference on Human Computer Interaction in Aeronautics (HCI-Aero'98)* (pp. 205-212), Montreal, Canada: Ecole Polytechnique de Montréal: 27-29 May 1998

Hollangell, E. (1993). The phenotype of erroneous actions. *International Journal of Man-Machine Studies, 39*(1), 1-32.

Hunt, R., & Rouse, W. (1981). Problem solving skills of maintenance trainees in diagnosing faults in simulated power plants. Human Factors, 23, 317-328.

Karsenty, L., Bigot, V., & de-Brito, G. (1995). *L'utilisation des procédures écrites dans un avion de nouvelle génération* (Report No. T-95-023). Toulouse, France: EURISCO Contactor Report.

Reason, J. (1990). *Human error*. Cambridge: Cambridge University Press.

Weiner, E.L., & Curry, R. (1980). Flight deck automation: promises and problems. *Ergonomics*, 23.

RÉSUMÉ

CONSCIENCE DE LA SITUATION ET SUIVI DE PROCÉDURES

Les procédures écrites utilisées en aéronautiques sont conçues comme une liste d'actions pour aider le pilote dans l'exécution des actions visant à garantir un niveau maximal de sécurité ainsi qu'une performance optimale. L'étude de l'utilisation des procédures opérationnelles Airbus a permis de mettre en évidence les raisons pour lesquelles les pilotes n'utilisent pas les procédures comme prescrit. Trois raisons principales sont avancées : tout d'abord, les pilotes doivent gérer des situations opérationnelles différentes de celles prévues dans les procédures ; de plus, suivre pas à pas une prescription ne leur permet de garder le contrôle de l'activité ; enfin, les pilotes ressentent le besoin constant de comprendre la situation dans laquelle ils se trouvent. Cet article suggère de nouvelles perspectives pour la conception des systèmes d'alarmes, pour la conception des procédures opérationnelles et pour la définition du rôle du pilote.

MOTS CLÉS : Résultats empiriques, Suivi de procédures, Conscience de la situation, Systèmes à risques

Human Machine Cooperation in the Anesthetic Consultation : Importance of Planning Activities for Information Gathering.

F. Anceaux*, M.C. Beuscart-Zéphir, P. Sockeel*****

*CNRS - UHVC, LAMIH- Equipe PERCOTEC, BP 311, F-59304 Valenciennes Cedex.
Francoise.Anceaux@univ-valenciennes.fr
**CERIM, Faculté de médecine, Université de Lille 1, Place Verdun, 59041 Lille Cedex, France
mcbeuscart@aol.com
***URECA, Université Charles de Gaulle, 59653 Villeneuve d'Ascq Cedex, France
sockeel@univ-lille3.fr

ABSTRACT

We present an evaluation of a software application supporting direct entry of data during the anesthetic consultation. The results demonstrate that this tool is inadequate at both on quantitative and qualitative level. These poor results are due to the incompatibility between the software and the anesthetist's usual strategies for their consultation. The software does not allow the anesthetist to rely on his expertise (i) to properly manage the dialog with the patient and the gathering of the essential medical data necessary for the planning of the anesthetic process, and (ii) to enter these data while they occur in the dialog.

KEYWORDS

Anesthesia, Planning, Information Gathering, Dynamic Process

INTRODUCTION

In the past fifteen years financial and legal problems in the healthcare domain as well as new medical trends have constituted strong motives to introduce software applications in the hospital setting. But as soon as these new tools deal with medical activity and imply physician-machine cooperation, their integration into the daily working environment often results in a failure.

Our study focuses on anesthetist's activity. We present an evaluation of a new software application supporting direct entry of data during the anesthetic consultation. This tool proved to be incompatible with the anesthetist's usual strategies for their consultation. From the analysis of this integration failure, we show the necessity to take into account, not only the user's requirements, but also his cognitive activity when prototyping a new application supporting medical activity (Beuscart-Zéphir, Brender, Beuscart, & Depriester, 1997 ; Beuscart-Zéphir, Sockeel, Bossard, & Beuscart, 1998).

BACKGROUND

Most of the studies on medical activity focus on reasoning, diagnosis, and therapeutic decisions (Boreham, Foster, & Mawer, 1992 ; Boreham, Mawer, & Foster, 1996 ; Groen & Patel, 1988 ; Patel & Groen, 1986). Recently, several authors suggested that medical activity or «healthcare» could be considered as the management of a dynamic process, such as survey and control of physiological processes (De Keyser & Nyssen, 1993 ; Hoc & Amalberti, 1995 ; Hoc, 1996). This kind of approaches highlights the processes underlying information selection and interpretation as well as the planning of activities involved in complex situations.

According to this approach, Xiao, Milgram, & Doyle (1997) identify four phases in the anesthetic process : pre-operative preparation, induction, maintenance and survey of the recovery phase. Until now, studies have focused on induction and , maintenance phases (Gaba, 1991 ; Gaba, Fish, & Howard, 1994 ; Gaba, Howard, & Small, 1995 ; Nyssen & Javaux, 1996).

The pre-operative phase is mainly devoted to the planning of the three remaining phases. The anesthetists try to anticipate the potential problems and to assess the risks of the anesthesia for the patient. In this respect, the anesthetist has to scan the patient's medical background exhaustively.

In France, the pre-operative phase is set on legal grounds. The patient has a medical consultation with an anesthetist one week before the actual surgical operation. The anesthetist who performs this consultation may not be the same person as the one actually in charge of the remaining phases. So the anesthetic file filled by the physician during the consultation is of major importance because:

- it is a legal medical record ;
- it has the important function of transmitting the relevant medical information to the anesthetist on duty during the surgery. Thus, it conveys both the underlying planning of the anesthetic process elaborated during the consultation, and the entire medical information necessary for the management of the induction and maintenance tasks. It allows the anesthetist in charge to assess those plans and to eventually modify them according to the evolution of the patient medical status during the surgery.

The medical data collected by the anesthetist are written down on a specific one-page sheet of paper. This paper file ordinarily contains 9 fields (administrative data, medical and surgical antecedents...) ; each field is divided into several zones, each one being devoted to one main physiological system. All along the interview of the patient and the clinical examination, the anesthetists fills in the given fields and zones. However, as requested by new legal dispositions, this anesthetic medical file must be typewritten.

For that purpose, a software application was developed to allow the anesthetist to enter typewritten medical data directly during the consultation. This tool, named AMS (standing for Anesthetic Mobile System), is a kind of computerized medical file, specific for the anesthetic consultation.

The AMS prototype was elaborated following a standard conception cycle, including an extensive User's Requirement phase (Brender, 1997) performed by both engineers and consultants of the project. Moreover, two expert anesthetists participated in the phase of elicitation of expertise, which was performed by an expert physician specialized in Hospital Information System databases. The main issue identified in this phase emphasizes a strong requirement for the exhaustiveness of the database. As most of the medical data are entered in the patient file through catalogs, the anesthetists want to be able to get from these catalogs **any** relevant medical item.

The general structure underlying the AMS interface is close to the one page paper sheet structure. But due to the exhaustiveness requirement, each field necessitates one screen page divided into several catalogs. Thus, the AMS interface is made up of eleven screen-pages divided into several fields, each field being devoted to the entry of a specific item. The logic of completion of the AMS sheet is therefore different from the logic of completion of the paper sheet : when he needs to jump from one field to another, the anesthetist has to shift the screen page, which implies a minimum number of actions.

Once the prototype had been achieved and technically verified, we were asked to assess its usability in a real work environment

MATERIAL AND METHODS

First step : Analysis of the Activity in Natural Setting.

Before introducing the new application, we performed an observation, description and analysis of the anesthetist's activity during consultation, with a particular focus on the interview of the patient and on data acquisition. The methods were: interviews, video recording of the consultations, auto-facing interviews.

The main objective of this analysis was to identify the procedures for searching, selecting and recording the relevant information. From the analysis of the patient/anesthetist dialogs and of the corresponding paper files, we built up, for each consultation, activity diagrams representing the order of the questions asked by the physician, the answers given by the patient, and the resulting written data.

Second step : Evaluation of the Usability of the Software Application.

Seven anesthetists participated in the evaluation study. Before the experiment, they were trained with the AMS application during three sessions of one hour.

Each anesthetist performed two consultations for one patient : one with the usual one-page sheet, the other with the AMS application. The order of the two consultations was randomized. Subjects were real patients, volunteering to participate in the study. Consultations were audio-recorded and the software application was equipped with a sneak system recording all the actions of the anesthetists. In the paper condition, two experimenters recorded on an observation grid the order and location of data acquisition. The method previously reported was used for the analysis of the data and the elaboration of the diagrams.

RESULTS AND INTERPRETATION

Analysis of Activity in Natural Setting

Results

All the patients are asked the same general set of questions, but the interview does not usually go through the successive fields in a systematic way. Nevertheless, the order of the questions differ from one case to the other. This order depends on two interdependent causes : the degree of complexity of the case, and the procedure for exploring the patient background.

From the analysis of the diagrams, we could identify three different procedures for the exploration of the patient's medical case. Each procedure accounts for an observed order of the questions. All the three procedures can be used alternately during the same examination. While the procedure 1 is prominently used in simple cases, complex cases show an alternate use of the three different procedures with procedures 2 (and 1) being more frequent. This mixing of the procedures leads the anesthetist to jump frequently from one field to another.

- *Procedure 1:* the anesthetist follows a standard and systematic order when questioning the patient, field by field and system by system.
- *Procedure 2:* from an answer given by the patient, the anesthetist infers some further relevant information and sets specific questions to confirm this hypothesis. This procedure leads to significant short cuts in the exploration of the patient's medical framework.
- *Procedure 3:* at times, the anesthetist may allow the patient to «tell his story» as far as it is relevant to the purpose of the consultation.

In these procedures 2 and 3, the order of the questions is no more the standard and systematic one.

Data are handwritten on the anesthetic consultation one-page paper-file. A lot of abbreviations are used to facilitate note taking ; important information and recommendation are underlined. The location of the data written on the paper file can be different for each anesthetist. Most of the data are written down as the corresponding information occurs in the dialog, leading the anesthetist to jump frequently from one field to another of the paper sheet.

Interpretation

The management of information gathering ,which expresses itself in the three above-mentioned procedures, seems to depends on three main sources :

- the procedure 1 is mainly supported by the structure of the anesthetic one page paper sheet ;
- the procedure 2 is driven by the anesthetist expert knowledge, which directs the interpretation and selection of relevant data, and the exploration of the patient's medical background ;
- the procedure 3 relies mainly on the answers of the patient, which provide raw information.

The alternate use of these three procedures constitutes a mixed bottom-up / top-down planning (Hoc, 1987), usually called opportunistic planning (Hayes-Roth & Hayes-Roth, 1979), which is partly supported by the actual acquisition of data. All along the examination, the anesthetist has the whole data already acquired before his eyes. As the examination progresses, the one-page file becomes a kind of external transcription of the temporary representation of the patient's medical framework supporting the dynamic planning of the medical examination.

Experimental Evaluation

Results

♦ Quantitative results.

In the AMS condition, the duration of the consultation is almost six times longer than in the « paper » situation.

Duration of consultation	Minimal	Maximal	Mean
Paper condition	5	10	7
AMS condition	30	60	40

Table 1: duration of the consultation (minutes)

The AMS record contains more data (10 to 20 %) than the paper record. But these supplementary data are found to be mostly administrative cues. Conversely, some essential medical data are missing in the AMS record.

♦ Qualitative results.

In the AMS situation, the anesthetist cannot perform the interview as in the paper situation because he is unable to jump as quickly from one field or location of the file to another as required by the rapidity of the dialog : too many operations must be carried out on the computer. Consequently, the dialogs show a lot of dead-ends and returns in the first 10 minutes. At the beginning, the anesthetist performs his interview as in a natural situation, trying to use short cuts or information spontaneously given by the patients. But he then finds himself unable to enter the corresponding data. After a few setbacks, the anesthetist quits those procedures 2 and 3, and follows the order of the questions set by the AMS interface. The interview is then driven by the structure of the interface, and the resulting dialogs look like some uneasy formal questioning. In addition, the anesthetist never gets in one glance the whole data already entered. Then, after a few minutes, he is unable to remember exactly which data have been acquired.

Interpretation.

Most of the increase in consultation time is due to the incompatibility between the characteristics of the interface and the way the anesthetists usually perform their consultation.

Dead-ends in the dialog and especially in data acquisition are mainly due to the difficulty to use procedures 2 and 3 with the AMS. This results for the anesthetist in the impossibility to apply a personal opportunistic planning.

Moreover, all the anesthetists claim to be «lost» after a few minutes. This confirms the importance of the one-page file for supporting the planning activity. In the AMS situation, the anesthetists are unable to embrace simultaneously all the relevant information. Then they seem to be unable to build up a proper representation of the patient's medical case.

This demonstrates a strong relation between the planning activity on the one hand, and the progressive elaboration of the representation of the patient's case on the other hand. The ultimate aim of the planning is to elaborate a representation of the patient's medical framework, which is significant for the whole anesthetic process ; each collected medical data is therefore relevant and makes sense because it belongs to this significant frame. Conversely, if this representation is insufficient or incomplete, the anesthetist cannot go on planning the exploration of the patient's background. He then has no choice but to follow the questioning order of the interface. Finally, without the support of the representation, the list of items and series of questions don't fully make sense, which explains why important information can be missing in the AMS record.

CONCLUSION

The consultation is an important phase of the whole anesthetic process mainly devoted to the assessment of risk and planning of the remaining phases. The anesthetic medical file is thus a very important document, because it conveys a representation of the patient's medical framework and contains all the information necessary to generate and adapt the plans of the anesthetic process.

The anesthetic consultation itself proves to be a dynamic activity aiming at the production of a significant and adequate representation of the patient's medical state, properly transcribed on the anesthetic medical file. These dynamic cognitive processes were not properly taken into account for the development of the AMS software application

The anesthetist's expert knowledge was correctly integrated in the software ; but the structure of the interface leads unavoidably to a stiffening of the procedures, making the anesthetist unable to implement any personal planning.

This failure is partly due to the strong user's requirement for exhaustiveness. The anesthetists require that this new tool be fully exhaustive, aiming at the «zero error» level. But the resulting tool prevents the anesthetists from relying on their own expertise to manage the risk. This proves to be potentially dangerous (Amalberti , 1996).

Further research is therefore necessary to conceive and elaborate such a software application as would allow a real man-machine cooperation and maintain a human management of the risk.

REFERENCES

Amalberti, R. (1996). *La conduite de systèmes à risques*. Paris : P.U.F. (collection Le Travail Humain).

Beuscart-Zéphir, M.C., Brender, J., Beuscart, R., & Ménager-Depriester, I. (1997). Cognitive evaluation : how to assess the usability of information technology in healthcare. *Computer Methods and Programs in Biomedicine, 54,* 19-28.

Beuscart-Zéphir, M.C., Sockeel, P., Bossard, B., & Beuscart, R. (1998). Activity modeling for assessing the usability of telematics applications in healthcare. In B. Cesnik & al (Eds.), *Medinfo'98, Amsterdam, IOS Press, IMIA*, 832-836.

Boreham, N.C., Foster, R.W., & Mawer, G.E. (1992). Strategies and knowledge in the control of the symptoms of a chronic illness. *Le travail humain, 55 (1)*, 15-34.

Boreham, N.C., Mawer, G.E., & Foster, R.W. (1996.) Medical diagnosis from circumstantial evidence. *Le travail humain, 59 (1)*, 69-86.

Brender, J. (1997). *Methodology for assessment of medical IT-Based Systems*. Amsterdam : IOS Press (Studies in health technology and informatics).De Keyser, V., & Nyssen, A.S. (1993). Les erreurs humaines en anesthésie. *Le Travail Humain, 56 (2-3)*, 243-266.

Gaba, D.M. (1991). Human performance issues in anesthesia patient safety. *Problems in anesthesia, 5*, 329-350.

Gaba, D.M., Fish, K., & Howard, S.K. (1994). *Crisis management in anesthesia*. New York : Churchill Livingstone.

Gaba, D.M., Howard, & S.K., Small, S.D. (1995). Situation awareness in anesthesiology, *Human Factors, 37(1)*, 20-3Groen, G.J., & Patel, V.L. (1988). The relationship between comprehension and reasoning in medical expertise. In M.T.H. Chi, R. Glaser, & M.J. Farr (Eds.), *The nature of expertise* (pp. 287-310). Hillsdale, N.J. : Erlbaum.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science, 3*, 275-310.

Hoc, J.M. (1987). *Psychologie cognitive de la planification*. Grenoble : P.U.G.

Hoc, J.M. (1996). *Supervision et contrôle de processus : la cognition en situation dynamique*. Grenoble : P.U.G.

Hoc, J.M., & Amalberti, R. (1995). Diagnosis : some theoretical questions raised by applied research. *CPC, 14 (1)*, 73-101.

Nyssen, A.S., & Javaux, D. (1996). Analysis of synchronization constraints and associated errors in collective work environments. *Ergonomics, 39 (10)*, 1249-1265.

Patel, V.L., & Groen, G.J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, *10*, 91-116.

Xiao, Y., Milgram, P., & Doyle, J.D. (1997). Planning behavior and its functional role in interactions with complex systems. *IEEE transactions on systems, man, and cybernetic - Part A : systems and human*, *27 (3)*, 313-324.

RESUME

Coopération homme-machine dans la consultation d'anesthésie : Importance des activités de planification dans la prise d'information.

Nous présentons l'évaluation d'un nouvel outil informatique de prise de notes permettant la saisie directe de données par les anesthésistes durant la consultation pré anesthésique. Nous analysons les dialogues patient / anesthésiste ainsi que les données saisies par l'anesthésiste et ce, que la prise de notes soit médiatisée par l'outil ou non. Les résultats montrent que cet outil n'est satisfaisant ni d'un point de vue quantitatif (temps de consultation fortement rallongé...), ni d'un point de vue qualitatif (oubli d'informations essentielles...). Il semble que l'outil, qui rigidifie la procédure de notation des informations, ne permette pas à l'anesthésiste de mettre en œuvre les différentes stratégies de recherche et de sélection de l'information qu'il utilise dans la situation naturelle. De ce fait, il ne peut pas s'appuyer sur son expertise pour (i) gérer le dialogue et la recherche d'informations et (ii) noter les informations essentielles à la planification de l'acte anesthésique au fur et à mesure qu'elle apparaissent dans le dialogue et l'examen clinique.

EVALUATING DESIGN OF HUMAN-MACHINE COOPERATION: THE COGNITIVE WALKTHROUGH FOR OPERATING PROCEDURES

David G. Novick and Meriem Chater

EURISCO
4 Avenue Edouard Belin
31400 Toulouse, France

LIHS
Université Toulouse I
31000 Toulouse, France

{novick, chater}@onecert.fr

ABSTRACT

The methodology of development of interfaces can be adapted to development of operating procedures. In particular, the cognitive walkthrough can be adapted to account for steps and resources outside the computer's part of the system interface. Empirical evaluation suggests that a cognitive walkthrough for operating procedures (CW-OP) is reasonably efficient and can provide useful information for developers.

Keywords

Methodology, operating procedures, cognitive walkthrough

INTRODUCTION

We seek to introduce methodological support for human-centered development of operating procedures for process control. Grudin (1990) that the user's interface to a computer is not just the hardware and software that compose the computer. From the user's view, the interface includes an array of associated elements in the context of use, including documentation, training, and advice from colleagues. This suggests that the methodology of development of interfaces could be transferred or adapted to development of operating procedures. In particular, we focus on the adaptation and use of the cognitive walkthrough (Wharton, Bradford, Jeffries, & Franzke, 1992) for the evaluation of operating procedures and their documentation. We ask two key questions: (1) Can the cognitive walkthrough for operating procedures be used effectively by evaluators? and (2) Does the cognitive walkthrough for operating procedures provide perspectives of value to developers of procedures and their documentation? To address these questions, we will review the use of and need for procedures in aviation, review current human-centered methodologies for development of procedures, introduce the cognitive walkthrough for procedures (CW-OP), and report empirical results from use of the walkthrough.

METHODOLOGY OF PROCEDURES

In some complex, dynamic, and risky domains, such as aviation, operating procedures are an inevitable part of the crew's interaction with the system. An operating procedure exists to specify, unambiguously, what the task is, when the task should be conducted, how the task should be done, by whom it should be conducted, and what feedback is provided to other agents (Degani & Wiener, 1997). To a certain extent, procedures exist to deal with irreducible issues of the design of interfaces; humans may have a better capacity to handle dynamic or other complex and difficult situations.

A number of approaches for development of operating procedures have been proposed in the field of aviation. While researchers generally agree that the field lacks a genuine methodology of procedure development, these approaches are steps forward. The most prominent approach is Degani and Wiener's (1997) "Four P's" model, which incorporates the organization's *philosophy* of operations, their business *policies*, *procedures* that effectuate the operations consistently with the policies, and the crews' actual *practices* on the flight deck. Within the design process itself, however, Degani and Wiener's model does not offer specific guidance on assessing usefulness or usability. Another approach that relates procedures to documentation is the act-function-phase (AFP) model (Novick & Tazi, 1998). While AFP provides a basis for classifying kinds of acts in operating procedures, it does not yet provide means of developing the procedures as such. A third approach was proposed by Drury and Rangel (1996) for reducing automation-related errors in aircraft maintenance and inspection. This approach does have a strong analytic component, which concentrates on identifying error types and opportunities in the procedures.

These approaches all encompass some degree of human-centered design, but tend to rely on post-design testing or interviews. What is missing is an *exploratory inspection method for evaluating use*, so that issues of both usefulness and usability (Gould & Lewis, 1983) could be addressed early in the design process.

THE CW-OP

Developers of user interfaces could address these issues through the use of a human-centered analytical technique such as the cognitive walkthrough (Lewis, Polson, Wharton, & Rieman, 1991; Wharton, Bradford, Jeffries, & Franzke, 1992; Wharton, Rieman, Lewis, & Polson, 1994). The cognitive walkthrough is a usability inspection method for interfaces that originally focused on evaluating a design for ease of learning. The method leads the designer to consider factors such as users' backgrounds and mental effort. It is based on an interaction model like that of Norman's (1986) stages of user activities. In brief, a task is decomposed into steps in the interface, which are analyzed individually with respect to connections between goals, artifacts, actions, and results. Each step is described as a "success story“ or a "failure story,“ depending on the outcome of the analysis. An extensive literature, summarised by Wharton, Bradford, Jeffries, and Franzke (1992), has examined the relative effectiveness of the cognitive walkthrough.

From the most practical version of cognitive walkthrough (Wharton, Rieman, Lewis, & Polson, 1994), we developed a new version adapted to operating procedures. This involved an iterative process of revision and use of forms and instructions adapted to procedures and their documentation. The CW-OP included five key changes:

- Dealing with procedural rather than interface steps. Some procedural steps may not involve action in the interface. This addresses the issue of including human-human as well as human-machine interaction.
- Drawing the evaluator's attention to the representation of the procedure in the documentation. This is the way the procedure exists in the context of use.
- Asking the evaluator to determine explicitly if training or experience are required for a particular step. These factors are typical justifications for a link between, for example, goals and actions. Requirement of training or experience may indicate that the procedure or the documentation should be modified.
- Looking at whether the procedure correctly implements the intended function with respect to the overall system. This addresses issues of usefulness and safety.
- Determining whether errors are probable and, if so, whether these would have an effect on safety. In the domain of flight operations, safety is the top priority.

The current version of the CW-OP, like the cognitive walkthrough for the physical interface, is supported by the use of two forms: (1) a cover sheet and (2) a form that presents the success or failure "story" each step analyzed. The contents of the cover sheet, which are basically the same as that in the original walkthrough, are presented in Figure 1. The contents of the form for an individual step are presented in Figure 2. The second form reflects the five changes adapting the walkthrough to procedures. In the actual forms, much more space is provided.

CW-OP Cover Sheet
Date:
Analyst(s):
Users:
Interface:
Task:
Action sequence:
Comments:

Figure 1. Contents of the Cover Sheet

CW-OP Story: ( ) *Success* ( ) *Failure*
Date:
Analyst:
Task:
Step:

Walkthrough

1. Will the users try to achieve the intended effect?
2. Will the users notice that the correct action is available?
    a. Documentation
    b. Interface
3. Will the users associate the correct action with the effect trying to be achieved?
4. If the correct action is performed, will the users see that progress is being made toward solution of the task?

Observations

1. Are experience or training needed?
    If so,
    a. Is this kind of step common or rare?
    b. Will training be easy or difficult?
2. Is the step correct in terms of function?
3. Are particular errors likely?
    If so, what is their impact on safety?
4. Design suggestions
5. Other comments

Figure 2. Contents of the procedure-step form

EVALUATION

An empirical evaluation of the cognitive walkthrough for procedures was conducted in order to address the study's main questions, which were operationalised in terms of the following hypotheses: (1) The CW-OP could incorporate elements dealing with procedures and documentation without undue burden on evaluators; (2) the CW-OP would identify issues involving the procedural as well as the physical interface; and (3) the evaluators' assessments would show a high level of agreement.

The hypotheses were tested through a walkthrough of draft operating procedures for a proposed text-based cockpit interface for air-traffic control (ATC) communications. Figure 3 presents one of the draft procedures. The test used six evaluators, including a computer scientist, two industrial ergonomists, two doctoral students in computer science and ergonomics, and one graduate-student intern in computer science. The evaluation was preceded a half-day of training. The evaluators were given three pages of a draft manual explaining the physical interface and the meaning of the air-traffic-control messages, along with three draft procedures. Taken together, the procedures encompassed eight unique top-level steps. Evaluators performed their own decomposition of steps into sub-steps as they judged appropriate. The tasks corresponded to the procedures. The action sequences that served as the standard for evaluation were determined by the procedures themselves plus the regulatory standards from which the interface and the procedures were developed.

Results

Following the CW-OP, all forms were reviewed for completeness, and the responses to individual questions aggregated.

Hypothesis 1

The data confirmed the hypothesis that the CW-OP could incorporate elements dealing with procedures and documentation without undue burden on evaluators. All evaluators completed the task within the 90-minute period. The number of steps analyzed per evaluator ranged between 8 and 13. These rates appear consistent with those reported for the non-procedure walkthrough (cf., Lewis, Polson, Wharton, & Rieman, 1991). The evaluators all expressed the opinion that the session had been valuable for them.

Hypothesis 2

The data confirmed the hypothesis that the CW-OP would highlight issues involving the procedures as well as the interface. As indicated in Table 1, 31 of 48 comments or design suggestions concerned procedures and their documentation rather than the physical interface.

Most of the step comments about the physical interface were actually questions about how the interface or system worked. The design suggestions included items such as changes to procedure embedding, wording, level of detail, and sequencing of steps. Additionally, there were four overall comments on the evaluation process itself, including suggestions for reworking the evaluation forms. The form in Figure 2 reflects some of these suggestions.

Whether the evaluators' findings were useful in the redesign of the procedures is a different question. While difficult to quantify, development following the evaluation included redesign of all three procedures that either eliminated or reformulated each procedure using specific findings reported on the forms.

Procedure: Respond to a clearance

- If you **need a short-term delay** to respond to the clearance, respond STBY.
- Do one of these cases:

  If you are **able to accept** the clearance, respond WILCO.

  If you are **unable to accept** the clearance, respond UNABLE.

Example

Aircraft receives message "AT ALCOA CLB TO & MAINT FL390."

Aircraft sends message "WILCO."

Figure 3. Draft procedure "Respond to a clearance"

Hypothesis 3

The data were inconclusive in confirming the hypothesis that the evaluators would agree in their assessments. The raw distributions suggest reasonable agreement. Application of Cohen's Kappa statistic (Carletta, 1996), which assesses whether classification by multiple coders exceeds chance levels, suggests that the agreement is better than chance but that confirmation of inter-rater reliability requires a larger test set. These results can be interpreted in terms of evaluators' responses to representative questions on the step form, including success/failure, availability of the action in the documentation, and availability of the action in the interface.

Success/Failure Stories. There was some disagreement among the evaluators as to whether the procedure steps presented success or failure stories. Raw percentage agreement among all six coders was 0.750. However, given that there were only two categories, percentage agreement was necessarily at least 0.500. Looking at Kappa, which ranges from 0 (no agreement) to 1 (perfect agreement), K=0.40 for classification of success/failure. This indicates that there was a good measure of agreement above chance but not as high as generally sought for classification of reliable categories. In fact, Kappa is not stable at low *n*, as is the case here. Consequently, these values are considered reasonable for exploratory work involving initial sessions by relatively untrained evaluators while coding categories are being developed.

	Physical interface	Procedures or doc
Overall comments	0%	2%
Step comments	28%	22%
Step design suggestions	6%	38%
Total	34%	62%

Table 1. Distribution of evaluator responses: Comments and suggestions (N = 48)

	Story for a Procedure Step?	
	Success	Failure
Step 1.1	CDEF	AB
Step 1.2	DE	ABCF
Step 2.1		ABCDEF
Step 2.2	ABCDEF	
Step 2.3	BCDE	AF
Step 2.4	ABCDE	F
Step 3.1	BCDE	AF
Step 3.2	BCE	ADF

Table 2. Distribution of evaluator responses:
Is the step a success or failure story?

A more informative view of data such as these comes from direct examination of the distribution of responses, which is presented in Table 2. The six evaluators are identified by letters that are consistent across the table. The eight steps are numbered for purposes of the table; these numbers were not in the materials provided to the evaluators. The table indicates strong agreement among the evaluators on three of the eight steps, and an even split on one step. This suggests that developers of operating procedures might benefit from understanding the reasons behind the evaluators' disagreements. For example, evaluator F considered that task sharing and coordination between the crew members was not sufficiently explicit.

Action Available in Documentation. The distribution of ratings on availability of the correct action in the documentation, presented in Table 3, suggests reasonably strong levels of agreement. Kappa was highly unstable on this distribution, and thus was not meaningful. The distribution suggests, though, that most evaluators will agree on whether the documentation provides the right action. For step 2.1, the evaluators generally found that an action term in the procedure was not explained in the documentation.

Action Available in Interface. The distribution of ratings on availability of the correct action in the interface, presented in Table 4, suggests moderate levels of agreement. Kappa was highly unstable on this distribution, and thus was not meaningful.

For step 2.1, the evaluators all found that an action term in the procedure was not a label in the interface. For step 2.3, some of the evaluators found that a label's meaning was confusing. For step 3.2, some evaluators found that the documentation referred only to the goal rather than the action.

	Action Available in Documentation?	
	Available	Not Available
Step 1.1	ABCDEF	
Step 1.2	ABCDEF	
Step 2.1	D	ABCEF
Step 2.2	ABCDEF	
Step 2.3	BCDEF	A
Step 2.4	ABCDEF	
Step 3.1	ABCDEF	
Step 3.2	BCDEF	A

Table 3. Distribution of evaluator responses:
Is the correct action available in the documentation?

	Action Available in Interface?	
	Available	Not Available
Step 1.1	BCDEF	A
Step 1.2	ABCDEF	
Step 2.1		ABCDEF
Step 2.2	ABCDEF	
Step 2.3	BCE	ADF
Step 2.4	ABCDEF	
Step 3.1	ABCEF	D
Step 3.2	BCE	ADF

Table 4. Distribution of evaluator responses:
Is the correct action available in the interface?

The range of evaluator agreement for these questions on the step form suggests that multiple evaluators can provide complementary perspectives on the usefulness and usability of operating procedures through use of the CW-OP. Multiple evaluators bring a variety of experiences with interfaces and procedures, different kinds of knowledge about human factors, and different levels of expertise about the interface being evaluated.

DISCUSSION

The empirical evaluation suggests that the cognitive walkthrough for procedures can be used effectively by evaluators and that it provides perspectives of value to developers of procedures and their documentation. Post-experiment debriefings and reviews of the study's evaluation forms suggested a number of improvements in the evaluation process, including training and coding.

The study involved parallel individual evaluations by persons with training in cognitive science. Users could be involved directly in the evaluation through use of the group-style cognitive walkthrough (Wharton, Rieman, Lewis, & Polson, 1994).

Training should include some individual (i.e., non-group evaluations) in order to make sure that each person has experience in answering all questions. For example, one evaluator reported difficulty in making the success/failure distinction.

Developers may want to ask evaluators to mark sections they read or used during the evaluation process. For example, no evaluator provided comments on the examples included as part of the procedures. Because there was no record of whether the evaluators used the examples to understand the procedures, the process did not provide a basis for assessing whether the examples were helpful.

Some other aspects of the procedures and their documentation did not get evaluated explicitly in the walkthrough. First, no one evaluated the names of the procedures, presumably because they are not steps as such. This leaves open the question of how to obtain analyses of the role of the name in accessing procedures as a part of their context of use. Second, overall instructions tend not to be evaluated. For example, the draft manual contained an instruction on when members of the crew should make announcements to each other during procedures. There was only one comment on this instruction in the session. How to obtain analyses of such overall instructions remains an open question.

Analysis of the evaluations and post-comments also suggests that the task model presented to evaluators was relatively weak. In determining the correct action during evaluation of steps in the computer interface, the evaluators used the interface description provided in the draft documentation, backed up by (1) an available reference document for the requirements set by the regulatory authorities and (2) their personal knowledge of the interface. The number of questions written by the evaluators with respect to the operation of the interface and its underlying system, as reported in Table 1, indicate that these resources were not sufficient.

This experience suggests using, to the extent possible, an explicit "gold-standard" definition of the interface that would eliminate uncertainty about correct actions in the physical interface and the functions of the underlying system. However, requiring a precisely defined physical interface makes it more difficult to develop the procedural interface while the physical interface has not been specified in detail. This may limit the method's usefulness during early development phases.

An alternative to the CW-OP might be the GOMS method (John, 1995), which enables describing the task and the user's knowledge needed to perform it in terms of goals, operators, methods, and selection rules. GOMS can predict skilled-performance time, method-learning time, operator sequence and likelihood of memory errors. GOMS can also be used to design training programs and help systems as it describes the content of task-oriented documentation. However, GOMS may not be appropriate for use early in the design process of operating procedures, particularly where issues of usability and safety are more important that issues of timing. GOMS is of relatively little help on design issues independent of the procedural quality of the interface including: (1) standard human factors issues such as visual quality of a display layout, readability of words and letters (2) the quality of work environment, user acceptance, and (3) the social and organizational impact of the system (John & Kieras, 1996). And, most important for operating procedures, GOMS does not help assess factors such as guidance and feedback.

Aside from the direct assessment of usability, the CW-OP provides additional insight into usefulness and safety. In particular, the cognitive walkthrough's requirement of a reason for linking goals to actions leads evaluators to determine whether training or experience are required for correct use of the procedure. Reductions in training required to operate an aircraft would both increase reliability and lower operating costs.

ACKNOWLEDGMENTS

This research was supported by a contract from Aerospatiale Aeronautique. We thank Florence Buratto, Laurent Moussault, Florence Reuzeau and Stéphane Sikorski for their participation in and contributions to the research, and we thank Jean Carletta for her advice about how to use and interpret findings on inter-rater reliability.

REFERENCES

Carletta, J. (1996). Assessing agreement on classification tasks: The Kappa statistic. *Computational Linguistics*, 22(2), 249-254.

Degani, A., and Wiener, E. Procedures in complex systems: The airline cockpit. *IEEE Transactions on Systems, Man, and Cybernetics*, 27, 3 (1997), 302-312.

Drury, C. G., & Rangel, J. (1996). Reducing automation-related errors in maintenance and inspection. In *Human Factors in Aviation Maintenance—Phase VI: Progress Report Vol. II* (pp. 281-306). Washington, DC: Federal Aviation Administration/Office of Aviation Medicine.

Gould, J., and Lewis, C. (1983). Designing for usability: Key principles and what designers think. *Proceedings of the Conference on Human Factors in Computing Systems (CHI 83)* (pp. 50-53). New York, NY: ACM Press.

Grudin, J. (1990). interface. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 90)* (pp. 269-278). New York, NY: ACM Press.

John, B. (1995). "Why GOMS?" *ACM Interactions*, 2(4), 81-89.

John, B., and Kieras, D. (1996). "Using GOMS for user face design and evaluation: Which technique?" *ACM Transactions on Computer-Human Interaction*, 3(4), 287-319.

Lewis, C., Polson, P., Wharton, C., and Rieman, J. (1991). Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. *Proceedings of the Conference on Human Factors in Computing Systems (CHI 91)* (pp. 235-242). New York, NY: ACM Press.

Norman, D. (1986). Cognitive engineering. In Norman, D., and Draper, S. (Eds.), *User centered system design*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Novick, D., and Tazi, S. (1998). Flight crew operating manuals as dialogue: The act-function-phase model. *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero'98)* (pp. 179-184). Montreal, Canada: Editions de l'Ecole Polytechnique de Montréal.

Wharton, C., Bradford, J., Jeffries, R., and Franzke, M. (1992). Applying cognitive walkthroughs to more complex user interfaces: Experiences, issues and recommendations. *Proceedings of the Conference on Human Factors in Computing Systems (CHI 92)* (pp. 381-388). New York, NY: ACM Press.

Wharton, C., Rieman, J., Lewis, C., and Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In Nielsen, J., and Mack, R. (Eds.), *Usability inspection methods*. New York, NY: John Wiley & Sons, Inc.

RÉSUMÉ

ÉVALUATION DE LA CONCEPTION DE LA COOPÉRATION HOMME-MACHINE : LA RÉPÉTITION COGNITIVE POUR LES PROCÉDURES OPÉRATOIRES

Notre objectif est de proposer un support méthodologique pour la conception centrée sur l'homme des procédures opérationnelles pour le contrôle des processus. Les méthodologies de développement d'interfaces homme-machine peuvent être adaptées aux procédures opérationnelles. C'est le cas de la méthode de la répétition cognitive qui a été adaptée à l'évaluation des procédures et de leur documentation. Cette adaptation porte sur cinq points : (a) prendre en compte les étapes de la répétition au niveau des procédures et non pas au niveau de l'interface ; (b) attirer l'attention de l'évaluateur sur la présentation des procédures dans la documentation ; (c) demander à l'évaluateur de déterminer explicitement si une formation ou une expérience est nécessaire ; (d) déterminer si la procédure met en œuvre correctement la fonction recherchée ; (e) déterminer la probabilité d'occurrence des erreurs et leurs implications pour la sécurité. L'évaluation empirique montre que cette méthode est efficace et peut fournir des informations utiles aux développeurs.

MOTS CLÉS : Méthodologie, Procédures opérationnelles, Répétition cognitive.

Session 2: Computer assistance
Les assistances informatiques

From analysis of watch officer's activity to ergonomic assessment of Automatic Radar Plotting Aid

Christine Chauvin
Laboratoire Ergonomie et Sécurité des Activités Maritimes,
Université Bretagne Sud, Centre de Recherche
Rue Saint-Maudé, 56325 Lorient
Tel : (33) 02 97 87 45 21, Fax : 02 97 87 45 00,
e-mail : christine.chauvin@univ-ubs.fr

ABSTRACT
This paper presents an analysis of watch officer's activity during collision avoidance. The analysis has two purposes. It must be of use to evaluate radar interface; it must be of help to define new functions or new aiding tools.

Key words: Navigation at Sea, Collision Avoidance, Human-machine co-operation.

INTRODUCTION

Two dangers threaten a boat: grounding and collision. The activity performed to avoid a collision consists in deciding on an action after having made a diagnosis of the situation. During this activity, watch officers use a main tool: the ARPA[1] radar. This tool has many functions, which make the activity easier but its usability and utility is sometimes criticised. Concerning usability, watch officers contest user-friendliness of some radar interfaces. However, opinion is divided and ships' managers are waiting for ergonomic criteria to be able to compare existing ARPA radars objectively. Concerning utility, system-makers consider that ARPA functions are not sufficient to prevent collisions.

To answer these requests and questions, an analysis of watch officers' activity was carried out. This paper consists of four parts. The first part is a presentation of the activity fulfilled during collision avoidance. The second part describes the method used to realise activity's analysis. The third part is a review of the main results. The fourth part provides some answers to the questions of usability and utility of radar functions and suggests means to improve the human-machine co-operation.

ACTIVITY OF WATCH OFFICER DURING COLLISION AVOIDANCE

Several researches deal with watch officers' behaviour, performance and reasoning during collision avoidance. They describe collision avoidance like a two stage process consisting of the diagnosis and the avoiding action. During this process, ARPA assists the mariner in resolving the problem of tracking targets and analysing their movements. ARPA is a computer *incorporated in the radar/ ARPA system so that the ARPA data can be displayed on the same screen as the conventional radar data* (Bole & Dineley, 1995). This system shows useful information concerning the target but does not give adequate judgement, decision-making and manoeuvre orders.

The diagnosis

The diagnosis itself can be divided into three steps: risk detection, risk evaluation, choice of manoeuvre.

To detect the risk before the introduction of radar, watch officers used a method related to the concept of 'steady bearing'. In fact, the risk of collision can be ascertained by carefully watching the compass bearing of an approaching vessel: *such risk shall be deemed to exist if the compass bearing of an approaching vessel does not appreciably change* (Cockroft & Lameijer, 1996). They took the bearing in observing the detected objects. Since the introduction of ARPA radar, new tools are available. The Electronic Bearing Line (EBL) enables the measurement of the bearing of vessels on the radar screen, by setting the line on the target and by observing whether the bearing changes over some minutes. Thanks to the 'plotting' function, watch officers may display courses and speeds of vessels in the form of vectors; they can conclude that a risk exists when the projected courses and speeds of two vessels place them at or near the same location simultaneously. Lastly, the watch officer may use the DCPA as a criterion for determining if the risk of collision exists. The DCPA is the Distance

---

[1] ARPA : Automatic Radar Plotting Aid

between two vessels at the Closest Point of Approach. If this distance is equal or near to 0 N.M.[2], then a risk of collision does exist.

Risk evaluation results from a comparison between the DCPA and a 'safety-distance', which is the range to keep clear in order to prevent any predictable collision accident. The amount of this range is related to the size, speed and manoeuvrability of the ship. It is defined by each watch officer and not by rules. Risk evaluation is therefore, the result of a subjective judgement.

When the DCPA is lower than the 'safety-distance', watch officers have to plan a manoeuvre according to the International Regulations for Preventing Collisions at Sea (HMSO, 1972). Collision regulations include three main rules defining three kinds of conflict: overtaking, head-on situation and crossing situation. In the case of overtaking, *any vessel overtaking any other shall keep out of the way of the vessel being overtaken* (Rule 13). *When two power-driven vessels are meeting on reciprocal or nearly reciprocal courses so as to involve risk of collision, each shall alter her course to starboard so that each shall pass on the port side of the other* (Rule 14). *When two power-driven vessels are crossing so as to involve risk of collision, the vessel which has the other on her own starboard side shall keep out of the way and shall, if the circumstances of the case admit, avoid crossing ahead of the other vessel* (Rule 15). Rules 13, 14 and 15 give recommendation about the direction of manoeuvre (starboard or port). Collision regulations provide also recommendation concerning time and amplitude of the course or speed alteration. It is said that *every vessel which is directed to keep out of the way of another vessel shall, so far as possible, take early and substantial action to keep well clear* (Rule 16). The most difficult case is encountered when the 'give-way vessel' does not alter her course. In this case the privileged vessel *may take action to avoid collision by her manoeuvre alone, as soon as it becomes apparent to her that the vessel required to keep out of the way is not taking appropriate action in compliance with these rules* (Rule 17).

The avoiding action

The avoiding action itself can be divided into three steps (Koyama & Yan, 1987): avoiding (i.e. changing course to achieve a sufficient miss distance), avoided (i.e. going steadily on the original course after avoiding) and returning (i.e. changing course to return to the original path when the risk is over).

Studies dealing with collision avoidance show that the possible interpretation of the Regulations generate a certain uncertainty concerning actions of vessels. Uncertainty concerns:

i) the evaluation of the risk made by watch officers (Hara, 1993). If they do not consider that a risk does exist, they will not alter the vessel course,

ii) the distance at which the burdened vessel will alter her course (Habberley & Taylor, 1989),

iii) the direction of the course alteration (Hinsch, 1996).

Uncertainty is one of the problems encountered in diagnosis activities and emphasised by Hoc (1991), Roth and Woods (1988). About collision avoidance at sea, it is assumed that the best way to reduce uncertainty and the resulting stress consists in performing the most 'typical' manoeuvre. The same result was found out concerning car driving (Saad, 1996).

METHOD

The aim of the study is to know, on one hand, which ARPA functions are used during collision avoidance, at which stage of the activity and for what kind of purpose. On the other hand, the purpose is to know what the difficulties encountered by watch officers during this activity are. These purposes guided the choice of field situations and the choice of three dependent variables; i.e.: actions on ARPA (showing use of ARPA functions), verbalisations and features of the manoeuvre. These data have been processed in order to point out regularities in actions executed and, as far as verbalisations are concerned, to indicate mental representations built by watch officers.

Two field situations

Analysis has been performed from ecological data, collected onboard two ferries during 19 voyages. Both ships do the same route: they cross the channel between Ouistreham and Portsmouth. But they are fitted with two different ARPA radars. Activity of watch officers has been observed when ships cross traffic bearing up for the TSD[3] of Dover-Strait and traffic bearing up for the TSD of Casquets. During this time, which lasts around one and a half hour, the vessels have often to manoeuvre.

Data

Everytime a conflict was detected, several data were recorded. There are independent variables and dependent ones. The latter reveal activity of the watch officer and make up a 'scenario'.

---

[2] N.M.: nautical mile

[3] TSD: Traffic Separation Device

Independent variables

They may have an effect on the activity. They are: *i)* the features of context (visibility, wind force and direction, current force and direction, traffic, advance or delay in regard to the Expected Time of Arrival), *ii)* the kind of conflict (meeting, crossing, overtaking), *iii)* the priority: is own ship burdened or privileged?

The value of each of these variables has been noted.

Dependent variables

They are behavioural data and verbal data.

- Watch officer behaviour is characterised by use of ARPA functions (actions on ARPA have been recorded by means of a video camera set in front the radar screen). It is characterised too by features of manoeuvre (angle of course alteration, distance from the other ship when the alteration is initiated, DCPA, distance to the original route); each alteration of course has been noted.
- Thinking-aloud protocols have been recorded because it is assumed that those data can be of help to identify reasoning and mental representations. Ericsson & Simon (1993) have shown that referents expressed could provide information about content of short-term memory. Some psycholinguistic researches (Caron-Pargue & Caron, 1989) showed that linguistic cues could provide knowledge about changes into representation (choice of verbs), about uncertainty (modal adverbs), decision-making (modal verbs) and reasoning (connectives).

Data processing

Field studies are characterised by four main features: complexity, temporal span, extension and cost. They make the statistical search for causal explanations difficult. Therefore, we just tried to describe regularities and correlations in data recorded: i.e. behavioural data and verbal data. Verbal data have been divided into utterances and coded according seven dimensions: the referents and six kind of linguistic cues (Chauvin, 1999). Two Factorial Correspondence Analyses have been carried out. The first one was executed with the aim of showing correlations between the referent expressed and the linguistic cues. The second one concerned just one referent (the ship manoeuvre) and resulted in pointing out correlations between linguistic cues (verbs chosen, modal verbs, modal adverb, and tenses...) and the current phase of watch officer activity (diagnosis and the different stages of the manoeuvre).

RESULTS

Results are descriptive. They have been translated into a model, so that the analysis can be validated.

Descriptive results

Analysis of behavioural data shows a typical chain of actions. After having plotted the target (function named ACQUIRE[4]), some officers carry out a 'steady bearing' by means of the EBL. All of them request alphanumeric data about the tracked target (function named DESIGN[5]). Those who prefer true vectors opt, afterwards, punctually for relative vectors (function named TRUE/REL[6]). Then appear course alterations (manoeuvre). Prior to the first one, watch officers may adjust the SCALE[7] (from 12 N.M. to 6 N.M.). Scenario ends as they CANCEL[8] the track and adjust once again the scale (to 12 N.M.). This chain of actions is made up of different states and of transitions, which determine succession of episodes. Events, which set off transitions, are the distance between vessel and target, as well as DCPA. Their value depends on the type of conflict and on the priority (burdened vessel or privileged one obliged to alter her course, according to rule 17).

Analysis of verbal data consisted, first of all, in identifying and categorising referents expressed by watch officers. Six categories have been found out, i.e.: 'own ship manoeuvre', 'target manoeuvre', 'features of conflict' (distance, DCPA...), 'features of vessels' (routes, speed), 'target kind', ' evaluations '. According to the model of Ericsson & Simon (Ibid.), these referents represent the content of the short-term memory. The second stage of the analysis consisted in identifying correlations between relevant linguistic cues (such as modal verbs, modal adverbs, and connectives...) and the referent expressed on one hand and between linguistic cues and

---

[4] ACQUIRE: initiate the tracking of a target. During the acquisition time, the system makes successive plots of the target's position in order to establish its speed and course. Its predicted movement is indicated by a vector.

[5] DESIGN: At the request of the observer, the following information is available in alphanumeric form in regard to any tracked target: present range of the target, present bearing of the target, DCPA – Distance at Closest Point of Approach, TCPA – Time to Closest Point of Approach, calculated time course of target, calculated true speed of target.

[6] TRUE/ REL: True vectors show proper course and speed of targets and own ship. Relative vectors show targets relative motion.

[7] SCALE: The ARPA facilities are available on several scales (from 1.5 N.M. to 24 N.M. inclusive). Range scale adjustments are made using the increase/ decrease function.

[8] CANCEL: cancel the track of an acquired target.

the current stage of the activity on the other hand. It provides some interesting results concerning the difficulties encountered by watch officers.

Modal adverbs and epistemic verbs are associated with two referents: 'target kind' and 'target manoeuvre'. So, it is possible to assume that watch officers cannot establish, with certainty, the target kind and cannot understand or anticipate, with certainty, the target manoeuvre.

' Own ship manoeuvre ' is connected to different arguments, which may indicate which information is processed to make the decision; namely, the relative speed of the target, the DCPA of targets, features of wind and current. When the connective is ' in order to ' or ' to ', arguments indicate the aim of the manoeuvre. In fact watch officers seem to assign two different functions to the manoeuvre: to move away from the target, but also to transmit a message to the officer onboard the target (like in this sentence: *I am going to alter the course a little to starboard, to induce him into moving*). It has been noticed, furthermore, that way of expressing ' own ship manoeuvre ' evolves throughout the activity. During the diagnosis stage, this referent is associated to modal verbs; the direction of manoeuvre (port or starboard) is chosen but nothing is said concerning its magnitude. Angle of the course alteration is expressed during the action stage. This characteristic reminds the 'under-specification' of action, which is considered as a cause of mistake (Reason, 1990). In the case of collision avoidance, 'under-specification' of action may oblige the watch officer to change his plan, when he has begun to alter the course. For example: he alters the course to starboard after having altered it to port, because he realises that it is not possible to cross ahead of the target. Such manoeuvres are dangerous because they are unusual and, as a consequence, difficult to understand or to anticipate by officers onboard potential targets.

Modelling

An object-oriented language (OMT), imported from computer sciences, was used for modelling (Rumbaugh, 1995). The model consists of three parts (Chauvin, 1996): a static, a dynamic and a functional one. They translate, respectively, the structure, the state and the working of the different 'objects' involved during collision avoidance, i.e.: ships, watch officers, environment, operating systems.

Static model

It shows hierarchical structure of the objects and presents it in terms of 'class'. It describes, particularly, the contents of the mental representation of watch officers, structure and the contents of their knowledge. These last are organised into two classes: knowledge of the conflict, knowledge of the ships, which divide in sub-classes depending on the type of conflict and type of ship. Knowledge of the conflict encloses data concerning the safety distance to keep between the ship and the target.

Dynamic model

It presents the different states of the objects. It shows, particularly, the different states in the reasoning of watch officers, ARPA functions used in each states and events, which set off the transformation from a state into an other. Reasoning is implemented to reach the following goals: 'identification of the class of conflict', 'evaluation of conflict', 'evaluation of bow or rear crossing range', 'choice of manoeuvre', 'identification of type of target', 'interpretation of present manoeuvre of target', 'anticipation of future manoeuvre of target', 'determination of magnitude of changing course to return to original path'.

Functional model

It deals with information processing and, therefore, with the reasoning performed by watch officers.

Twelve scenarios have been used to establish the results and the model. Data of eight other scenarios have been processed to valid it. The validation process confirmed the main results and enriched certain aspects of the model.

DISCUSSION

Thanks to the results of the activity analysis, it is possible to answer the requests and questions concerning 'usability' and 'utility' of ARPA functions.

Usability of ARPA functions

Results, translated by the static and dynamic models point out, on one hand, the main radar functions and, on the other hand, the sequence existing in the use of these functions. These results allow to apply to ARPA radar some concrete ergonomic recommendations concerning usability; and, particularly the following ones (Shneiderman, 1998).

- *Minimal input actions by user*. This rule leads to privilege direct manipulation rather than menus... at least for the most frequently used functions of the ARPA radar. Menus are, moreover, not convenient because there are few sequences in actions performed.

- *Prevent error.* Risk of error does exist when one control is used to alter both the bearing of the EBL and the range of the VRM[9]. In this case, it is necessary, in order to realise a 'steady bearing', to press ENTER to fix the EBL at its present position. If watch officers forget to press ENTER and want to use the VRM, EBL moves and 'steady bearing' has to be done once again. So, interface with two rotative controls (one to set the EBL and the other to set the VRM) seems preferable to interface with a common control.

Utility of ARPA functions

Results from verbal protocols analysis – translated into the functional model - can be use to uncover cognitive demands of the task, to identify errors likely to occur. Therefore, it can lead to the identification of the limits of the officer-ARPA system, to discuss utility of some aiding functions and to suggest the implementation of new ones.

Utility of the some of the aiding functions

It is interesting to search into the reasons why some aiding functions are not used, particularly in the case of the 'Trial Manoeuvre' function and of the 'DCPA and TCPA limits'.

- 'Trial Manoeuvre' function

The 'trial manoeuvre' feature allows the operator to see the results of possible changes in own ship's speed and/ or course, without actually implementing those changes. It could be useful to validate a choice if the computer would integrate the time of the course alteration; that is to say: if it could show what the situation would be if the changes occurred in 10 and 15 minutes.

- 'DCPA and TCPA limits'

Watch officers might enter the required DCPA and TCPA limits. If a target infringes both limits, an alarm is raised. In fact, they do not use this function. The reason could be that they use different limits, depending on the type of conflicts.

Towards most suitable tools...

Analysis shows that watch officers encounter two kinds of difficulties. The first one is related to data which remain uncertain (the target type and the target manoeuvre). Introduction of transponders onboard ships would solve this problem, since this tool provides precise information concerning the target and could provide information concerning its intention (Crichton & Redfern, 1996). The second difficulty concerns the planning of the manoeuvre. The most difficult aspect is not to define its magnitude, but to choose its direction (starboard or port side). As it has been pointed out in other diagnosis activities, problems encountered are related to the number of data to process (wind force, sea state, target speed, traffic, DCPA of a second target) and to their nature (changing and/or not well known). So, it is to support this reasoning that an aiding tool would prove most useful.

CONCLUSION

The analysis of watch officer's behaviour indicates what the functions of ARPA radar used during collision avoidance are. Analysis of thinking-aloud protocols reveals difficulties encountered to build facts (namely the target kind and target manoeuvre) and to make a proper decision about the direction of course alteration.

The first result leads us to advise to provide direct controls for the main functions and to make easier use of EBL. It gives, therefore, some concrete recommendations that permit the comparison of different interfaces.

The second one leads to the emphasis of the need for a device, which would provide information about target kind and target intention. Such a device - the transponder - does exist but is not installed onboard merchant ships. It leads also to emphasise the need for a device, which could help watch officers to plan a 'typical' (predictable) manoeuvre. Such a device would have to process several data. It requires, therefore, the integration of different tools (radar but also anemometer...), which is not realised yet. This device would be complementary to the transponder, which could, as a consequence, transmit reliable information about the action planned.

BIBLIOGRAPHY

Bole A.G., & Dineley W.O. (Eds.). (1995). *Radar and ARPA manual.* Oxford: Butterworth-Heinemann Ltd.

Caron-Pargue, J., & Caron, J. (1989). Processus psycholinguistiques et analyse des verbalisations dans une tâche cognitive. *Archives de Psychologie, 57*, 3-32.

Chauvin, C. (1996). *L'anticollision à bord de navires transbordeurs - Analyse et modélisation de l'activité des officiers de quart.* Thèse, Université René-Descartes, Paris.

Chauvin, C. (1999). Analyse de l'activité d'anticollision à bord des navires de commerce : des marques linguistiques aux représentations mentales. To be published in *Le Travail Humain.*

---

[9] VRM : the Variable Range Marker consists of a circle superimposed on the picture, centred on own ship's origin and adjustable in radius from zero to 99.9 N.M.. It allows rapid range measurements.

Cockcroft, A.N., & Lameijer, J. (Eds.). (1996). *A guide to the collision avoidance rules.* Oxford: Butterworth-Heinemann Ltd.
Crichton, T., & Redfern, A. (1996). Transponders and their impact on the future of navigation. *The Journal of Navigation, 49-3,* 332-336.
Ericsson, K.A., & Simon, H.A. (Eds.). (1993*). Protocol analysis - Verbal Reports as Data* (2ème édition). Cambridge: MIT Press.
Habberley, J.S., & Taylor, D.H. (1989). Simulated collision avoidance manoeuvres: a parametric study. *The Journal of Navigation, 42-2,* 248-254.
Hara, K. (1993). *A safe way of collision avoidance manoeuver based on manoeuvring standard using fuzzy reasoning model.* Paper presented at the *MARSIM'93 - International Conference on Marine Simulation and Ship Manoeuvrability.* St. John's, Newfoundland, Canada, Sept.
Hinsch, W. (1996). *Traffic rules to coordinate collision avoidance manoeuvres at sea.* Paper presented at the *International Conference on Preventing Collision at Sea – Collision'96.* Dalian, China, Sept.
HMSO (1972). *International Regulations for Preventing Collisions at Sea.* HMSO.
Hoc, J.M. (1991). *Some dimensions of a cognitive typology of process control situations.* Paper presented at the *Cognitive processes in complex tasks, 2nd Workshop.* Eindhoven, Pays-Bas, April.
Koyama, T., & Yan, J. (1987). *An expert system approach to collision avoidance.* Paper presented at the *8 th Ship Control Systems Symposium.* La Haye, Pays-Bas, Oct.
Reason, J. (1990). *Human error.* Cambridge: Cambridge University Press.
Roth, E.M., & Woods, D. (1988). Aiding human performance, I: cognitive analysis. *Le Travail Humain, 51,* 39-64.
Rumbaugh, J & al. (1995). *Modélisation et conception orientées objet.* Paris: Masson.
Saad, F. (1996). Driver strategies in car-following situations. In A.G. Gale (Ed.), *Vision in vehicles 5* (pp. 61-70). Amsterdam: Elsevier Science / North Holland.
Shneiderman, B. (Ed.). (1998). *Designing the user interface.* Addison Wesley Longman.

RÉSUMÉ

DE L'ANALYSE DE L'ACTIVITÉ DE L'OFFICIER DE QUART À L'ÉVALUATION ERGONOMIQUE DE L'AIDE AUTOMATIQUE AU POINTAGE RADAR

Dans des zones où le trafic est dense, les navires effectuent fréquemment des manœuvres pour éviter les collisions. Lors de l'activité d'anticollision, les officiers en passerelle utilisent le radar ARPA. Cet instrument les aide à détecter les "cibles", leur fournit des informations permettant d'évaluer le risque et de prendre la décision la plus appropriée concernant la manœuvre à effectuer. Cet appareil fait, cependant, l'objet de critiques portant sur son "utilisabilité" et son "utilité". Une analyse de l'activité a été réalisée à bord de deux car-ferries et à l'occasion de 19 traversées France-Angleterre. Elle a porté sur l'utilisation des fonctions du radar, les caractéristiques des manœuvres effectuées et sur les verbalisations des officiers. Elle a permis de mettre en évidence les fonctions du radar les plus souvent employées, mais aussi les difficultés rencontrées par les officiers pour mener à bien leur activité. Elle a conduit à produire des recommandations ergonomiques concernant l'interface radar, à proposer des hypothèses visant à expliquer les raisons de la non-utilisation de certaines fonctions et à suggérer de nouveaux outils susceptibles d'améliorer l'aide apportée par le radar ARPA.

MOTS CLÉS : Navigation en mer, Évitement de collision, Coopération homme-machine

Elastic interfaces. Maritime instrumentation as an example

Peter Bøgh Andersen

Center for Human Computer Interaction/Department of Information and Media Science, University of Aarhus. E-mail: pba@imv.aau.dk

ABSTRACT

This paper introduces the general notion of *elastic interfaces* and the derived concepts of *transparency* and *tailorability*, and exemplifies the notions in the domain of maritime instrumentation. It concludes that maritime automation to some degree already follows principles of transparency and possibly might benefit from adopting a disciplined version of tailorability.

Keywords: maritime instruments, tailorability, transparency, HCI.

ELASTIC INTERFACES

This paper reports on some ongoing work in the project *Elastic Systems* under the *Danish Center for Human Computer Interaction*[1].

First a definition: an *elastic interface* is one that lets the user move continuously in two dimensions:

1. The dimension of *access*. The dimension measures the degree to which the user has access to changing the underlying features of the system that determine its possibilities of use. At one extreme, the user is confined to only using the system for the purpose it was designed for. At the other extreme, the user has the power of a normal programming environment. The interesting part is the middle area of this dimension.
2. The dimension of *agency*. The dimension measures the degree to which the user or the system determines the course of events. At one end, the user initiates every action and the system merely responds. At the other extreme, the system handles the whole process and the user can only start and stop it. Again it is the middle part that is interesting.

The term *elastic* can be seen as a name for an already existing tendency, noted in Gentner & Nielsen 1996, towards designing systems that are *mutable* (buzz word: *tailorability*) and which may exhibit variable degrees of *autonomy* (buzz-word: *autonomous agents*).

The term *elastic* stresses that computer systems are dynamic systems that involve two main kinds of dynamics: a synchronic dynamics between user and system and a diachronic dynamics of adaptation and development. It invites us to view a computer system more as an evolving process than as a stable object. This paper applies the concept to the maritime domain that is characterized by two interesting features:

- The domain is a safety-critical domain which creates special conditions for tailoring.
- Maritime instruments by nature have an autonomous environment. The navigator cannot escape sharing control over instruments with wind, waves, and currents.

A simple maritime example of (1) is that some maritime systems can be toggled between sea and harbor conditions, displaying different types of information and enabling different actions in the two conditions. Alarms are a simple example of (2). It is a problem that maritime systems give too many and too undifferentiated alarms which the navigator must acknowledge and thereby possibly may get diverted from the real task at hand. Alarms are sounded on the initiative of the machine to call the attention of the operator and thus regulate the balance of power between the two.

In the domain of process control, the *access* to changing the system should be strongly constrained by safety considerations. Movement on the *agency* scale is already carefully regulated in the maritime domain.

THE CONCEPT OF TAILORING

Basically, tailoring is motivated by the inability of the designer to accurately predict the needs of the use situation, and most modern system therefore offer some kind of tailoring: setting preferences and redesigning the menu structure. End-user programming in a scripting language is not uncommon. However, there is often too large a gap between the understanding developed during usage and the concepts needed for tailoring purposes, and users are

---

[1] The center is funded by the Danish National Research Foundation. The empirical data are from three ferries, a freighter, a 2 weeks cruise with a large container ship, and from two transcribed simulated voyages at the Danish Maritime Institute. It must be borne in mind that the ship types are not yet representative.

often reluctant to accept this additional burden (Wasserschaff & Bentley 1997: 307). In Andersen 1997[1990]: 180 I suggested the following principle for tailorable systems:

1. Experiences gained from using the system should be applicable for modifying and changing it.

and later research in tailorability has emphasized the same point (Malone, T. W. K-Y. Lai & C. Fry. 1995: 178).

In this paper I shall explore the following very simple and general principle:

2. The geometry of tailorable artifacts should be self-similar.

By self-similarity I mean the mathematical property that parts of the object are a scaled down version of the whole object (see Peitgen, Jürgens & Saupe 1992 for a good treatment of the concept and many examples). Self-similar objects have the property that if a part is enlarged it displays the same morphology as the whole object. Applied to computer systems it means that the internal structure of a module of the system is similar to the structure of the whole system. This enables experiences suitable for handling the whole system to be applicable for understanding the details of the parts too. Two specific criteria can be deduced from this. One pertains to understanding and one to acting:

3. *The Principle of Transparency:* What goes on between two internal objects of the system is analogous to what goes on between the user and the interface-objects.
4. *The Principle of Tailorability:* Changing internal objects of the system is like changing the interface-objects of the system.

WHY TRANSPARENCY AND TAILORABILITY?

In this section I shall present two arguments for transparency and tailorability in systems for maritime navigation. The transparency argument concerns situations of malfunction of machinery, whereas the tailorability argument takes its point of departure in the fact that navigation is divided into phases and therefore causes the information needs to change.

Transparency: Handling and Preventing Malfunctions

When I visited a modern ferry with integrated bridge and many useful systems for aiding the captain, the two navigators pulled me aside when I was leaving: they were quite satisfied with their equipment but they wanted to voice one worry: *how do we know what is left when a component falls out?*

With the old equipment it was easy to figure out what was left if, for example, the GPS (Global Positioning System) stopped functioning. The GPS was an isolated instrument that showed the ship's position, but had no interconnections with the other instruments. So the captain knew "what was left" when the GPS malfunctioned.

This is much more difficult to assess in integrated bridges, where components deliver data to many other components. Thus, it is important that the user has a chance to *understand* the nature of malfunctions and figure out how the modules interact with each other.

Even more important is the ability of *acting* in the case of malfunction. If the automatic steering system breaks down, the navigator should be able to easily revert to manual steering, because the crew cannot stop work because of an instrument failure, as the user of a safe PC can do.

Prevention of malfunctions is another good reason for transparency. For example, in Maersk 1999 it is required that instruments and steering gear are regularly checked for correctness and functionality. If such control is to be possible at all, the workings of sensors and actuators must be understandable.

But creating such understanding is no trivial problem: if the automatic systems are "strong, silent, clumsy, and difficult to direct" (Woods 1996: 6) they generate confusion and errors.

Tailorability: Phases

A main argument for tailorability is the fact that navigation falls into distinct phases, each with their own information needs.

There are three main phases: navigation in open waters, coastal waters and harbor[2]. The phases are clearly marked by different types of cooperation and information needs. For example, in open waters one officer can handle the navigation, whereas two to four may be needed in the two other cases.

In one ship I visited there was the following sequence that reflects a decreasing curve of tension: (1) When things are difficult inside the harbor, the captain handles the steering advised by the pilot and helped by the chief officer. (2) Then the helmsman is called in and receives commands from the captain. (3) Outside the trafficked

---

[2] In production plants, phases can be found e.g. in the start-up procedure.

area, the pilot leaves, and the watch officer replaces the chief officer. (4) The helmsman leaves, and (5) in open waters (in the daytime and in clear weather) only the watch officer is on the bridge. The sequence 1⇒ 5 is used when leaving the harbor, whereas the reverse 5 ⇒ 1 is used during entry.

Not only the work organization, but also the place of work changes with the phases. During berthing, navigation can be moved to the bridge wing so that the navigator can see the ship's side. Each bridge wing must therefore contain a copy of the relevant displays and controls.

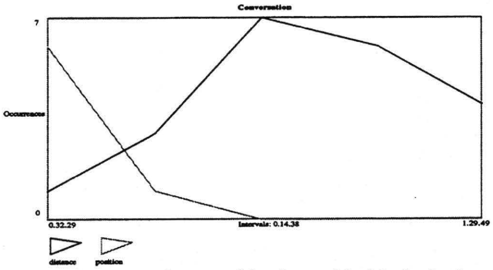


Fig. 1. Distance and not position is used inside the harbor.

Information needs change as well. Outside the harbor, the radar is used to plot bearings and distances to other ships, but inside there may be too many vessels too close for this to be useful. In open waters, the GPS information of latitude and longitude is useful for determining the position of the ship, whereas bearings to landmarks or buoys are used in coastal waters. Inside a harbor, the importance of visual assessments of distances increases; some instruments increase their relevance as well, while others change their function. For example, drift and set become important during berthing. The radar is now used for measuring the distance to the quay, and the V(oyage) M(anagement) S(ystem) for assessing the position of the ship in the basin. Other instruments become obsolete: for example engine information about revolutions per minute (rpm)[3], fins, and autopilot have nothing to do in the bridge wing that is specialized for berthing maneuvers. This is illustrated in Fig. 1. that displays the occurrences of conversations dealing with distance and positions from a simulated voyage at the Danish Maritime Institute (see Andersen 1998b).

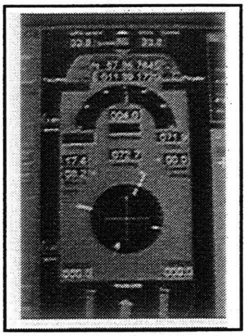


Fig. 2. Conning display

With a metaphor, one may say that while navigating a ship is like driving a train in open sea, it is more like riding a horse inside the harbor.

These shifts of information needs are not reflected in instrumentation. For example, for some strange reason, many instruments give latitude and longitude a prominent place in the display. In the conning display in Fig. 2, it occupies the "best" place in the top half of the display. However, the vessel was a ferry that always sailed in restricted waters and the numbers were never used (May 1999).

In cases like this, one might suspect that some kind of tailoring of the interface would be appropriate, allowing the captain to select the information relevant for the phase he is in and give it a prominent place.

This would not only benefit the person using the instruments, but also the other officers he is cooperating with since good bridge practice requires all to verbalize their actions so that everybody knows what goes on. This ability is taught on courses, is encouraged in ship owner's guidelines (Maersk 1999) and can in fact be observed in good captains:

(H, verbalizing)	I am down on ninety nine, a hundred now, ninety nine, or ninety eight
(E, volunteering information)	We are to start the turn when we have the buoy to the north

For more examples, see Andersen 1998a, Section 2. This kind of discrete monitoring and helping out is not unique to maritime navigation but is probably common in all kinds of process control, cf. Heath & Luff 1992 that conclude that in order for such mutual monitoring to occur, the size and location of the display is crucial.

There is therefore reason to believe that giving the phase-relevant information a prominent position on the bridge may enhance the mutual monitoring and help.

AGENCY AND ACCESS IN EXISTING TOOLS

In this section we shall see how the needs for transparency and tailorability are realized in existing instrumentation.

---

[3] Although the exact rpm number is irrelevant, there is still a need for a more coarse engine feed back telling the captain whether his command is being executed: does the ship move in the right direction and do the revolutions go towards what he has ordered (e.g. full ahead).

Agency

Agency is regulated in two ways, namely

- *Paradigmatic* (either-or): Either the system or the user is doing the task, but not both at the same time. The system take over tasks which the navigator will perform under other circumstances.
- *Syntagmatic* (both-and): Both system and user are active in performing the task, the system allowing the navigator to manipulate work objects of different granularity.

Voyage Management Systems exemplify the *paradigmatic* type: they allow the navigator to draw a track the ship should follow, and when a plan is executing, the system replaces the navigator as the controller of the rudder.

Integrated joysticks are examples of *syntagmatic* automation. On one ship, a joystick controlled the water-jets of the ship in unison, so that the navigator did not have to think about manipulating the individual jet, but could imagine that he manipulated the whole ship. The automation did not replace the navigator, but allowed him to work on a higher level of granularity.

Because of its safety critical nature, maritime instrumentation is based on the principle that if higher level automation falls out, there must always be a lower level that can be used. Steering is a good example. In one ship there were at least six following levels: VMS, Autopilot, Helm, N(on)F(ollow)U(p), Manual steering of rudder machine, Physical movement of rudder by means of wires (Figs. 3-5). Similarly, instead of the integrated joystick the navigator could use individual sticks controlling the individual jets, and he could even manipulate their hydraulics directly.

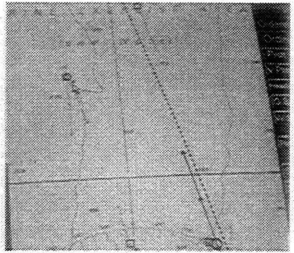


Fig. 3. VMS system on bridge

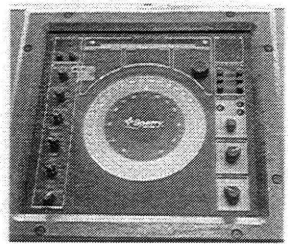


Fig. 4. Autopilot on bridge

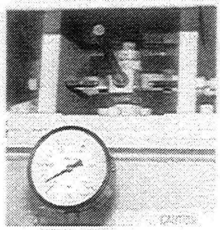


Fig. 5. Manual steering of rudder machine in engine room.

Shifts of agency is carefully regulated in both instrumentation and operating guidelines.

Access

Whereas shifts of agency play a major role in ship design, shifts of *access* are not so pronounced. However, electronic charts can be tailored, so that only a selection of the available buoys, beacons, lights, depth contours, cautionary areas, and currents are displayed. The purpose is to avoid cluttering the chart. Another area is configuring instruments.

USING TRANSPARENCY AND TAILORABILITY IN DESIGN OF MARITIME INSTRUMENTS

In this section I shall present some ideas for exploiting transparency and tailorability more systematically in the maritime domain.

Transparency

Some existing systems tend to conform partially to the principle of transparency. The higher level automation handles the lower level in the same way as the user does. For example, both VMS and user can input course orders to the autopilot, both autopilot and user can input rudder orders to the rudder system, and both rudder system and user can control the hydraulics of the rudder machine. Thus, the system is self-similar with respect to *pragmatics*, i.e. the user-system relation. I can understand the workings of the VMS in analogy to the way I myself interact with the autopilot.

One could consider extending the pragmatic self-similarity to *semantic* self-similarity. This would mean that the internal system structure is based on the concepts the navigator uses when manipulating the whole system. The purpose is to make it easier for the user to move along the dimension of agency. Using action-based intentional concepts for process control has already been suggested in Lind 1990, 1994. In order to achieve semantic self-similarity, the action schemata must be extracted from the language pertaining to the usage of the instruments, and in the maritime domain this means *maneuvers*. The following schema is extracted from an analysis of typical maneuvers (Andersen 1998a).

1. *Goal:* what we want to achieve, e.g. avoid collision with foreign ship.
2. *Means:* actions that lead to the goal, e.g. using the rudder.
3. *Help:* forces that enhance the means, e.g. exploiting the wind.
4. *Obstacles:* forces that hinder the goal, e.g. low speed.
5. *Negative side-effects:* negative state caused by the means, e.g. the turning moment produced during a crash stop achieved by putting the engine in reverse.
6. *Countermeasures:* actions that aim at eliminating obstacles and negative side-effects, e.g. compensating with rudder in (5) or increasing speed in (4).

For a system to be semantically self-similar its internal architecture must be based on concepts like these. If this is the case, the navigator can use his skills from maneuvering for understanding and repairing the system.

Tailorability

We cannot tailor collective, safety critical systems like a ship bridge in the same way as we do with our personal PC. Tailoring must be controlled, since it will not do that an officer has to look for the echo sounder in a new place each time he takes over the watch. Still, fixed and hardwired instruments do present problems for a work domain that is divided into phases. What is useful in one phase, is irrelevant for another. Furthermore, a panel that gives longitude and latitude a prominent position is well suited if the vessel is ocean-going, but becomes irrelevant if it is turned into a ferry.

Here is a possible scenario: sensor signals are dissociated from individual displays, and the hardwired displays are replaced by a panel of screens that can be treated logically as one screen so that is possible to move interface objects from one screen to another. Some interface objects, such as the VMS and the radar, are produced by manufacturers, whereas others objects can be designed by ship-owner, captain, or officers according to a well-defined hierarchy of constraints. For example, maritime law will rule that some instruments are obligatory and must have a certain size.

Judging from observations, the envisioned tool should offer facilities such as: moving and scaling objects, combining objects; changing mode of display. I conclude by two authentic examples that illustrate this.

1. An officer of the watch had forgotten that the rudder limit and R(ate) O(f) T(urn) limit were set and could not figure out why the wheel did not produce the desired turning. The reason was that the ROT display and the limits were represented in different places. In some autopilots limits are knobs whose position is difficult to see in the dark (see Fig. 4).

After having heard the story, another officer produced on the spot Fig. 7 as a solution to the problem. In his redesign he combined four old displays into one: the rudder angle and ROT indicators, and the two knobs from the autopilot. The curved bars indicate the angles of the rudder and rate of turn, and the bullets the limits of both. By looking at the new display the officer of the watch can immediately see the correct cause of the failure.

To realize this design, the officer needs facilities for changing the representation from the original digits to bars, and for combining four stand-alone instruments into one.

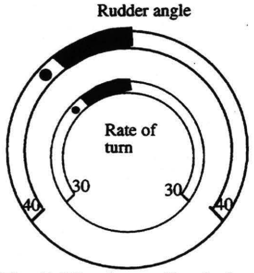


Fig. 7. "On the spot" redesign of displays

2. A captain wants to use fuel most economically, i.e. to achieve most speed for the lowest fuel consumption. On the Maihak display in front of him (Fig. 8) he sees 3 numbers on a display: revolutions per minute (rpm), torque, and horse power.

These numbers express the power produced by the engine. The captain wants to hit the most economical way of producing speed, so he wants to compare the three numbers with the speed indicator placed some distance away.

This is difficult two reasons: he must remember combinations of the four figures to hit upon the most economical one and his eye must move from one place in the panel to another one. Therefore, instead of looking at the numbers in front of him, he turns and follows the development of the rpm and speed on trend curves available on a screen behind him.

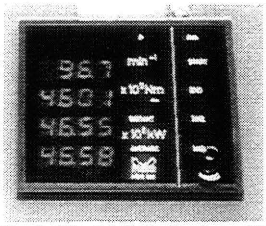


Fig. 8. The Maihak display with rpm, torque and horse power.

The curves are better in two respects: the rpm and speed curves are drawn in the same diagram and therefore it is easy to compare them and see when they increase in parallel and when speed starts to increase less than rpm. In addition, he does not need to remember past values, because this is what trend curves do.

To do this in reality the captain needs facilities for changing numbers into trend curves and combining two displays (rpm and speed) into one.

The examples are simple in the sense that they can be realized by changing representations and combining displays. However, a tailorable bridge could also stimulate imagination of the officers and give rise to new displays that the landlubbers of the instrument manufacturers would never think of.

REFERENCES

ANDERSEN, P. BØGH (1997 [1990]). A theory of computer semiotics. Semiotic approaches to construction and assessment of computer systems. Doctoral dissertation. Cambridge University Press.

ANDERSEN, P. BØGH (1998a). *Analysis of Maritime Operations II.* Center for Human-Machine Interaction. Report CHMI-2-1998. http://www.chmi.dk/

ANDERSEN, P. BØGH (1998b). *Analysis of Maritime Operations III.* Center for Human-Machine Interaction. Report CHMI-3-1998. http://www.chmi.dk/

CARROLL, J. & M. ROSSON (1987). Paradox of the Active User. In J. Carroll (ed.): *Interfacing Thought.* Cambridge, MA: MIT Press, 80-111.

DMI (1997). *Manoeuver Theory.*

GENTNER, D. & J. NIELSEN (1996). The Anti-Mac Interface. *CACM* 8: 70 - 82.

HEATH, C. & P. LUFF (1992). Collaboration and control. Crisis management and multimedia technology in London underground line control rooms. *Computer Supported Work* 1: 69-94.

HUTCHINS, E. (1996). *Cognition in the wild.* Cambridge, Mass.: The MIT Press)

LIND, M. (1990). *Representing goals and functions of complex systems. Institute of Automatic Control Systems.* DTU. 90-D-381.

LIND, M. (1994). Modeling goals and functions of complex industrial plants. *Applied Artificial Intelligence 8: 259-283.*

MACKAY, W. (1991). Triggers and Barriers to Customizing Software. In *Proc. of CHI'91,* New Orleans, Louisiana, 1991. New Work: The ACM Press, 153-160.

MALONE, T. W. K-Y. LAI & C. FRY. (1995) Experiments with Oval: A Radically Tailorable Tool for Cooperative Work. *ACM Transactions on Information Systems,* 13/2: 177-205.

MAY, M. (1999). Cognitive aspects of interface design and human-centered automation on the ship bridge: the example of ARPA/ECDIS integration. Presentation for PIC'99, Bath, June 21.23.

MAERSK (1999). *Guidelines for navigators. Maersk Vessels.* Technical organisation, level 2 document. NAU610. Version April, 1999

PEITGEN, H.O, H. JÜRGENS & D. SAUPE (1992). *Chaos and fractals.* Berlin: Springer-Verlag.

WASSERSCHAFF, M. & R. BENTLEY (1997). Supporting Cooperation through Customisation: The Tviews Approach. *Computer Supported Cooperative Work* 6: 305-325.

WOODS, D. D. (1996). Decomposing automation: apparent simplicity, real complexity. In: R. Paraduramen & M. Mouloua (eds.): *Automation and human performance: theory and applications.*, 1-17. Mahwah, New Jersey: Erlbaum.

RÉSUMÉ

LES INTERFACES ÉLASTIQUES : L'EXEMPLE DE L'INSTRUMENTATION MARITIME

Cette communication introduit la notion générale d'*interfaces élastiques* et les concepts dérivés de *transparence* et d'*adaptabilité*. Elle illustre ces notions dans le domaine de l'instrumentation maritime. Elle conclut que l'automation maritime suit déjà, dans une certaine mesure, les principes de transparence et qu'elle pourrait probablement tirer bénéfice de l'adoption d'une version modérée de l'adaptabilité.

MOTS CLÉS : Instrumentation maritime, Adaptabilité, Transparence, Interaction homme-ordinateur

SUPERVISED CO-OPERATIVE PATH PLANNER

Christophe LE ROY & Patrick CHEDMAIL

Equipe CMAO et Productique.
IRCyN. UMR C.N.R.S. 6597 - Ecole Centrale de Nantes.
1, rue de la Noë - BP 92101 - 44321 Nantes Cedex 3 - France
E-mail : Christophe.Le-Roy@ircyn.ec-nantes.fr

ABSTRACT

The validation of the mounting/dismounting simulation in a cluttered environment is a key problem during the design process of a mechanical system. On the one hand, research in path planning lead to automatic trajectory definition. On the other hand, direct manipulation is possible thanks to common virtual reality tools that allow the designer immersion in a whole mechanical environment. Thanks to the use of a multi-agent architecture we greatly improve the effectiveness of a path planning system coupling algorithmic approaches and human direct manipulation. Moreover, we manage to optimise parameters in order to simplify human integration and contribution.

Keywords : Virtual Reality, path planner, human integration.

1 INTRODUCTION.

For twenty years path planners for mobile objects in cluttered environments have been developed in the framework of robotics. Meanwhile, virtual reality (VR) tools are more and more used for industrial purpose. Such VR softwares enable the use of large virtual digital mock-ups to check the construction of mechanical structures through simulations. In order to help the designer, one of the main point is to simulate mounting/dismounting sequences between the different mechanical parts. For establishing these sequences it is compulsory to simulate object moving in such a way that these moves are out of collision. Our purpose is to enhance object manipulation and path planners with human abilities. Such abilities allow a global vision of the environment. Here after we propose to build a co-operation between an operator and path planner algorithms. This co-operation is built with multi-agent principles. We focus our analysis on the behaviour of the designer for finding best co-operation parameters.

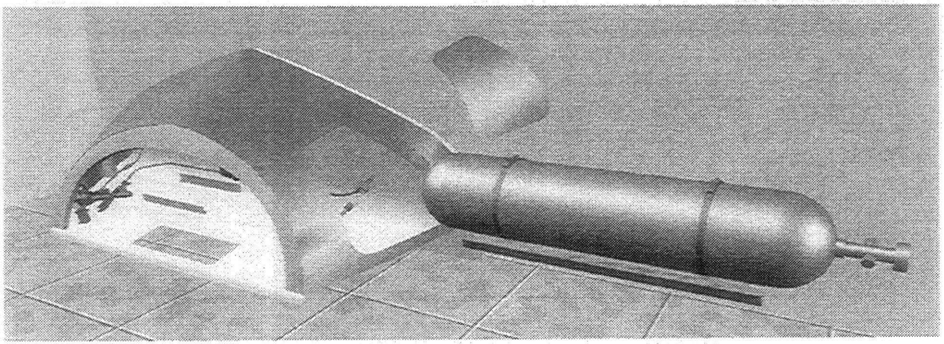


Figure 1 : Saab 2000 fore fuselage with the oxygen bottle.

2 STATE OF THE ART FOR PATH PLANNERS.

2.1 Classical approaches.

Many research topics deal with the definition of the collision free trajectories for solid objects. Moves considered are position, orientation or other degrees of freedom of kinematics. These moves are usually considered in the configurations space. Most of path planners are based on a partition dividing of the configurations space. Their heuristic usually ensures that a solution is found if there is no time constraint and if a solution exists. One of the main drawbacks is the time of research that grows in an exponential manner according to the number of degrees of freedom (dof). Moreover, the computation time increases with the state of clutter of the environment. Nevertheless, such path planners are really useful for simple problems.

In order to build a path planner, different approaches can be used. On the one hand some methodologies need a global perception of the environment. They imply a complete description of the configurations space. On the other hand, methodologies consider the moves of the object only in the close and local environment. One of the global techniques which uses the partition dividing of the configurations space is the Voronoï's method (Boissonnat, Faverjon, & Merlet, 1988). Barraquand and Latombe (1991) built another widespread global method. This method consists in a partition dividing of the configurations space and the use of potential fields in order to provide guidance for the trajectory search for a 2D robot. Local techniques are also quite numerous. The success of these methods is not guaranteed due to the existence of local minima. Khatib (1986) proposed a specific method based on potential field functions, as in Barraquand and Latombe method (1991). All these techniques are limited either by the computation cost or the existence of local minima.

2.2 Approach with human integration.

The use of path planner is acceptable for easy cases, without a lot of dof. For specific cases, as for numerous dof or for highly cluttered environments, it can be useful to be helped by human abilities. His or her abilities provide a global view of the environment. Consequently the path planning problem can be tackled in a more constructive way rather than if all is done automatically. Such kind of approach has been used by Hwang et al. (1997). Within Hwang's method, the designer can define a sequence of sub-goals in the environment. These sub-goals lead to the partitioning of the problem in several easy path planning problems. Different experiments with this sequential technique have shown the important potential of this method. This method confirms that human abilities enhance greatly path planners, even if human's intervention is sequential.

The points above point out clearly local abilities of several path planners. Furthermore, human global vision can lead to a coherent partitioning of the main trajectory. We intend to manage simultaneously these local and global abilities. The main purpose is to know how to build an interaction between human and algorithms in order to have an efficient path planner.

3 MULTI-AGENT APPROACH.

3.1 Previous studies.

Several studies about co-operation between processes have shown the great potential of co-operation between agents. These studies use multi-agent principles. First concepts were proposed by Ferber (1995). Latter on several studies checking co-operation between machines, humans and machines and between humans were performed. In the mechanical framework, some studies performed by Blanco, Garro, Brissaud, and Jeantet (1996) led to skill partitioning during the design process, over done by Chedmail, Damay, and Rouchon (1996), or Chedmail, Damay, and Yannou (1998) have shown creation of agents for geometric design of mechanical parts. These studies lead to the creation of a "Concurrent Engineering" methodology based on network principles, interacting with cells or modules that represent skills, rules or workgroups. Such studies can be linked to a work done by Arcand and Pelletier (1995) for the construction of a cognition based multi-agent architecture. This work presents a multi-agent architecture of human or society behaviour using cognitive psychology results within a human computer system.

These different studies show the important potential of Multi-Agent Systems (MAS). Consequently, we have built a path planner based on MAS with an interactive human integration within the processes rather than a sequential interaction as presented by Hwang (1997).

3.2 Elementary agent definition.

Several workgroup have established rules for the definition of an agent and its interactions, indeed even for moving architecture according to the environment evolution. Within these research groups it is important to point out contribution from Ferber (1995) and Jennings, Sycara, and Wooldridge (1998). From these analysis we consider next points for an elementary agent definition.

An agent :

1. is able to act in a common environment,
2. is driven by a set of tendencies (goal, satisfaction function,...),
3. has its own resources,
4. can see locally its environment,
5. has a partial representation of the environment,
6. has some skills and offers some services,
7. has a behaviour that intends to satisfy its goal, taking into account its resources and abilities, according to its environment analysis and to the information it receives.

The here-above points show that direct communications between agents are not considered.

3.3 Multi-agent architecture.

MAS theories allow different kinds of agent architecture. One kind of architecture is based on direct message communication between agents. In fact, our approach is based on the virtual environment, it implies that each agent acts on its set of variables from the environment according to its goal. Consequently, direct communication between agents is not compulsory. For these reasons an architecture based on a "blackboard" principle can fit our needs (fig. 2).

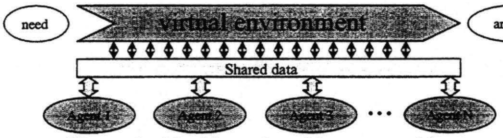


Figure 2 : blackboard multi-agent architecture based on the virtual environment.

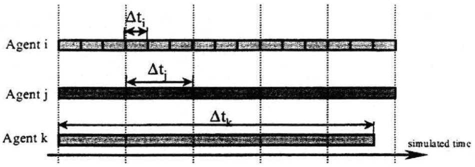


Figure 3 : agent activity rates.

Furthermore, our architecture is frozen, and the agents can not be cloned. However, their activity can evolve according to the task they have to achieve, each agent has a specific rate of activity $\lambda \in \mathbb{N}^*$. For example, figure 3 shows three agents. The rate of activity of agents i, j and k are respectively $\lambda_i = 1$, $\lambda_j = 3$ and $\lambda_k = 14$. For i and k, it means that agent i acts 14 more times than agent k. Moreover, any agent action on the environment is normalised. Besides human interactions with the environment, a designer can act on the different rates of activity of each agent (see next chapter).

4 EXPERIMENTS.

4.1 Multi-agent structure for the experiments.

Principles exposed in the previous chapter were used for our multi-agent architecture. Consequently for our path planner we have defined some local elementary agents which co-operate with a global human agent. This MAS interacts by means of shared variables which represent the environment status. Moreover, this architecture enables changes on the agent activity rates.

4.1.1 Elementary agents.

The first goal of the path planner is to reach a target. Consequently we consider the elementary agent *attraction*. The manipulated object is affected by an elementary move toward the target. This *attraction* agent only considers the target and does not take care of the environment.

The manipulated object must move without collision. The second elementary agent is a *repulsion* agent. This *repulsion* agent acts in order to nullify the collisions between the manipulated object and the cluttered environment.

For object build with several mobile parts, another agent is used. This *kinematic* agent takes into account the dof of the kinematic chain. The link between two parts can be rotational or prismatic. As for *repulsion*, the *kinematic* agent acts on the dof to avoid collisions with the environment.

Activity rates for the *attraction*, *repulsion* and *kinematic* agents are respectively λ_{att}, λ_{rep} and λ_{kin}.

4.1.2 *Operator agent.*

For our study it is compulsory to integrate a human operator within the multi-agent system. The *operator* agent does not fit exactly the elementary agent definition considered in 3.2. Points 4 and 5 are fortunately not taken into account. The *operator* has a global view of the cluttered environment displayed by the way of a virtual reality tool. Furthermore, *operator*'s action must be simple and efficient. So we use a Space Mouse which allows object manipulation with six degrees of freedom. The rate of activity for the *operator* agent is λ_{op}.

4.1.3 *Master viewer.*

As explain in point 3.3, a *master viewer* can control the rates of activity λ of the different agents. Any variation leads to another behaviour of the path planner. Consequently, this *master viewer* can get used with the system. He or she can act advisedly on agent activity rates. Our purpose is to analyse different samplings of rate of activity in order to define areas where our path planner is more efficient.

Final multi-agent architecture is schematised hereunder (fig.4).

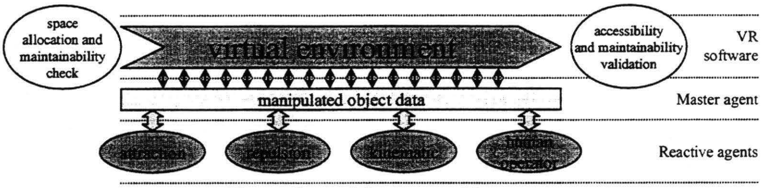


Figure 4 : multi-agent architecture.

4.2 Study on an industrial purpose.

4.2.1 *Main objective of industrial partners.*

After first studies on a simplified database we have tested our path planner on an industrial case. The initial purpose was the definition of accessibility sequences for mounting/dismounting simulations with a digital mock-up. Several live demonstrations have validated the manipulation of an oxygen bottle in the fore fuselage of a Saab 2000 aeroplane (fig.1). These experiments, which were performed with a human *operator*, have demonstrated the high adding value of human abilities.

4.2.2 *Modelled operator.*

For our industrial study we have local minima that can be avoided thanks to the *operator*. In order to build a criterion to define the best time sampling it was not possible to use a human operator for two main reasons. First reason is that a human does not always behave in the same manner on the manipulated object. The other reason is that the *operator* learns about the global behaviour of the MAS. As a consequence he can predict the local minima and the effectiveness of the path planner changes.

This *modelled operator* agent intends to attract the manipulated object toward a line that represents good position and orientation for the oxygen bottle to go through the door (fig.3). This agent works till the bottle has reached a quite good position and orientation in front of the door entrance.

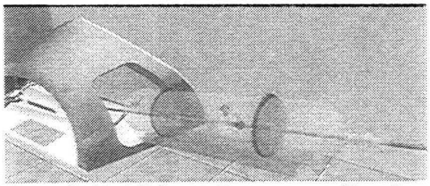


Figure 5 : Modelled operator represented by a line.

4.2.3 *Sampling variations.*

With the *modelled operator* we can be sure that our MAS is deterministic. Consequently, we can change agent activity rates in order to define best organisations. For this purpose we have built surfaces from a performance criterion applied to our path planner. Such kind of surfaces were used by Chedmail, Damay, and Yannou (1998) to find out best agents sharing out in order to build a mechanism.

4.3 Optimisation results.

Here we will consider two studies with different elementary contributions for the agents. The two studies A and B only differ for one parameter (normalised elementary movement for position). Results for studies A or B show that *repulsion* activity must be higher than *attraction* or *modelled operator* activities. Otherwise, in some cases the manipulated object passes through the cluttered environment, and that leads to a long time for trajectory research.

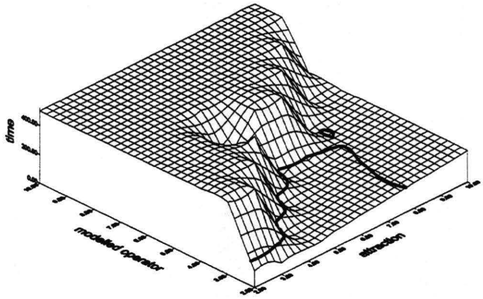


Figure 6 : surface from activity rate samplings for study A.

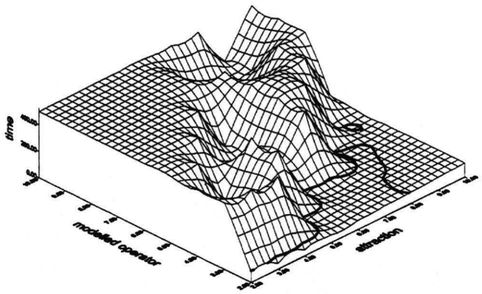


Figure 7 : surface from activity rate samplings for study B.

For study A, figure 6, we draw a line on the surface where the time of research is above 170 seconds. For study B, figure 7, the line corresponds to a time of 85 seconds. In each study the line defines a boarder that represents areas where time for trajectory research is under 170 seconds for study A, and 85 seconds for study B. Thank to these representations we have determined areas where our path planner is effective. The main role of the master viewer is to keep activity rates in these defined areas. Thus, the *master viewer* can help the *operator* agent providing him or her best coupling parameters for his or her objective. Further studies have shown that the *attraction* activity added to the *operator* activity must be under the *repulsion* activity. Moreover, we have verified that a reduction of the *attraction* activity when *operator* acts lead to a better co-operation. Over empirical experiments have shown that the best rate for the *kinematic* agent is $\lambda_{kin} = 1$ or 2. In this case the *kinematic* agent has a preponderant activity throughout the process for trajectory definition.

Thus, we must have : $\begin{cases} \lambda_{att} + \lambda_{op} < \lambda_{rep} \\ \lambda_{att} + \lambda_{op} \geq 8 \\ \lambda_{rep} \in [1;2] \\ \lambda_{kin} \in [1;2] \end{cases}$.

Finally we manage to build a path planner ; which is reactive to the *operator* needs, which is controlled by a *master viewer*. Moreover, the *master viewer* is over controlled thank to experiment results in order to assert the accuracy of our path planner architecture.

5 CONCLUSIONS.

Our study has shown that our MAS based path planner allowed a spatial manipulation of a kinematic object. Our system manages to tackle many degrees of freedom in a cluttered environment. Moreover, we have managed the integration of an *operator* in the multi-agent system that acts advisedly to help the system to reach an aim. First results encourage us to pursue with our specific architecture. Other studies have shown that the *master viewer* also acts advisedly to change the agent activity rates. If the *master viewer* controls activities according to the representations from point 4.3, we can be sure that our object manipulation will be performed without collision. Moreover, we can assert that for most of cases, if a trajectory exists for a path planning problem it can be found thank to the *operator* global view.

Furthermore, surfaces (fig. 6 and 7) can be used for the construction of an automatic "*master viewer agent*" that could act on the other agent activity rates. At this level we could compare our architecture to the one used by Arcand and Pelletier (1995) for their cognition based multi-agent architecture.

6 ACKNOWLEDGEMENTS.

This study has been supported by the European community, in the framework of the aeronautic CEDIX project (1998).

7 REFERENCES.

Arcand, J.-F., and Pelletier, S.-J., (1995), "Cognition Based Multi-Agent Architecture", *Intelligent Agent II, Agent Theories, Architectures and Languages*, August 19-20, 1995.

Barraquand, J., and Latombe, J.C., (1991), "Robot Motion Planning : A Distributed Representation Approach", *The International Journal of Robotics research*, December 1991.

Boissonnat, J.D., Faverjon, B., and Merlet, J.P., (1988), "Perception et planification", *Traité des nouvelles technologies*, Série Robotique.Hermès- Paris, France, 1988.

CEDIX : European community CEDIX project, "Concurrent Engineering based "design for integrated X-methodologies", no. BRPR-CT96-0151, BE95-1461", Synthesis Report, October, 20th, 1998.

Chedmail, P., Damay, T., and Rouchon, C., (1996), "Integrated design and artificial reality : accessibility and path planning", *ASME Engineering System Design and Analysis (ESDA'96)*, vol. 8, pp. 149-156, ISBN 0-7918-1503-X, July 1-4, 1996, Montpellier.

Chedmail, P., Damay, T., and Yannou, B., (1998), "A cooperative approach for dimensional synthesis of mechanisms", *Integrated Design and Process Technology, Society for Design and Process Science*, Berlin, Germany, July 6-9, 1998.

Ferber, J., (1995), "Les systèmes multi-agents, vers une intelligence collective", InterEditions, 1995.

Garro, O., Blanco, E., Brissaud, D. and Jeantet, A., (1996), "Intermediary object in the context of Distributed Design", *CESA'96, IEEE-SMC*, Lille, July 96.

Hwang, Y. K., Cho, K. R., Lee, S., Park, S. M., and Kang, S., (1997), "Human Computer Cooperation in Interactive Motion Planning", *ICAR'97*, Monterey, CA, July 7-9, pp. 571-576, 1997.

Jennings, N. R., Sycara, K. and Wooldridge, M., (1998), "Roadmap of agent research and development", *Autonomous agents and multi-agent systems*, Vol. 1, pp. 7-38, 1998.

Khatib, O., (1986), "Real time obstacle avoidance for manipulators and mobile robots", *International Journal Robotics Research*, Vol. 5, no. 1, pp.90-98, 1986.

Lozano-Pérez, T., and Wesley, M. A., (1979), "An algorithm for planning collision-free paths among polyhedral obstacles", Comm. ACM, 22 (10), pp. 560-570, 1979.

Tournassoud, P., (1990), "Géométrie et intelligence artificielle pour les robots", Hermès, Paris, France, 1990.

RÉSUMÉ

PLANIFICATEUR DE TRAJECTOIRE COOPÉRATIF SUPERVISÉ

La validation de la simulation de montage/démontage dans un environnement encombré est un problème clé au cours du processus de conception d'un système mécanique. D'un côté, la recherche en planification de trajectoire conduit à la définition automatique de la trajectoire. De l'autre, la manipulation directe est possible, grâce à des outils communs de réalité virtuelle qui permettent l'immersion du concepteur dans un environnement mécanique complet. Grâce à l'utilisation d'une architecture multi-agent, nous avons considérablement amélioré l'efficacité d'un système de planification de trajectoire en couplant des approches algorithmiques et la manipulation directe par l'opérateur humain. De plus, nous gérons des optimisations de paramètres pour simplifier l'intégration et la contribution humaines.

MOTS CLÉS : Réalité virtuelle, Planificateur de trajectoire, Intégration de l'homme

Adapting operational experiences to unexpected situations

Charlotte Skourup* and Arthur B. Aune**

*ABB Corporate Research
**Department of Engineering Cybernetics
Norwegian University of Science and Technology
charlotte.skourup@no.abb.com, aba@itk.ntnu.no

ABSTRACT

This paper outlines an experience repository that aims to manage human observations obtained from the process as well as intuition and knowledge concerning unexpected situations. In the real world, introducing one operator support system more in the control room often involves unforeseen problems. The proposed framework utilises the strengths of the computer and the human operator to share the tasks of the operator support system. Since knowledge-based behaviour is a very demanding process, it is desirable to support the operator at this level of behaviour. Human information processing in real-time tasks are extremely situation- and person-dependent.

Keywords

Operator support, task allocation, plant operation, visual decision support.

INTRODUCTION

Until today, much research has been directed towards improving the man-machine interaction and in particular, information transfer between the operator and the system whereas the information and knowledge transfer among the human operators has been attached less importance. The work presented in this paper introduces an operator support system based on a common experience repository for process plants, that stores information and knowledge, with a focus on implicit knowledge held by operators about unexpected situations. The main purpose of this framework is to share individual experience between the operating staff members. Moreover, the strengths of computers and human beings have been utilised to share the responsibility for information processing to obtain an optimal interaction, also among the humans.

Operators acquire knowledge through work experience and special training sessions. There exist two types of knowledge; explicit knowledge and implicit knowledge. The latter type of knowledge is often individual and hence, difficult to share with the other members of ones shift and particularly, with the rest of the operating staff. However, it is natural for humans to use previous experience in the problem solving process (Kolodner 1993, Rasmussen 1983), but, from an organisational viewpoint, individual knowledge is generally not fully utilised (Nonaka 1994).

Existing operator support systems normally provide support for operators in critical situations that can be predicted by, for example, mock-ups and mathematical models based on the first principles. But these support systems may in general have difficulties in analysing and classifying unexpected situations and hence, in suggesting useful solutions.

EXPERIENCE MANAGING IN PROCESS PLANTS

Process plants generally run 24 hours a day. Operators working in shifts are dependent on transferring information and knowledge specifically during shift exchange, but also between operators at different skill levels. A main part of this work includes visits to four major process plants in Norway and Sweden; the latter one is Hydro Polymers AB. This plant is a chemical process plant producing polyvinyl chloride (PVC). The main purpose of these visits has been to become familiar with operators' work tasks and experience and to get feedback on new solutions. Experience from the visits has mainly been based on observations of and interviews with the operators and the domain experts such as process engineers. Specifically, suggestions for improvements as well as feedback on the proposed framework for this operator support system, has been the main motivation for maintaining an effective dialogue with operators and domain experts. In particular, at Hydro Polymers AB, they have introduced a case-based reasoning system to administer experience concerning problem situations. The plant incorporate five shifts with twelve operators on each. The operators rotate the work tasks such that everybody performs all sort of tasks. Hence, the frequency of accomplishing certain tasks may be up to three months. Even then, it is not sure that the operator will meet any problems and thus, acquire experience concerning running this particular section. Generally, a problem does not necessarily need to be critical to proceed valuable experience. In critical situations, the operator has to intervene in the process to stabilise the situation whether the operator knows which actions to perform, or not. The work presented in this paper focuses on unexpected situations. Such situations are not necessarily critical. In this work, an unexpected situation is defined as a problem situation in which the operator does not have a ready response immediately. Often, the management underestimates unexpected situations, as they do not represent a direct risk for loss in profit or for the safety of the employees and the equipment.

AN EXPERIENCE-BASED FRAMEWORK FOR UNEXPECTED SITUATIONS

This paper outlines an experience repository that aims to manage human observations obtained from the process as well as intuition and knowledge concerning unexpected situations (Skourup and Alty 1998). Human beings very often reason by retrieving experience that are, in some sense, similar to a new problem and by reusing such experience to solve this problem (Rasmussen 1983). The case-based reasoning paradigm is a technique based on the very same, simple idea which can handle incomplete, uncertain and inconsistent domains (Kolodner 1993). The framework presented in this paper uses case-based reasoning to structure information and knowledge, and to search within the experience repository which all the operating staff members have access to. Kitano and Shimazu (1996) have presented a successful example of implementing case-based reasoning within an organisation. In the framework proposed in this paper, operators and domain experts such as process engineers collaborate on creating new cases. The operator defines new unexpected situations as tentative cases based on personal observations and individual knowledge such as the causes, actions to perform and outcome of these actions. This situation description can be extended at any time. As domain experts evaluate the tentative cases, they change the status from tentative to real cases. Domain experts may also perform additional analyses of each case. Thus, the quality of the cases is secured such that the experience repository only includes essential cases representing non-overlapping prototypes of problem situations. In addition, the domain experts may also define new cases. Both operators and experts incorporate individual knowledge in the cases reflecting situation descriptions based on their personal perception of the situation and its context as well as knowledge and intuition regarding the features such as the importance of each feature in proportion to the situation context. The proposed procedure for defining new cases is equal to the procedure which operators and domain experts in Hydro Polymers AB use for building up the content of an effective experience repository.

In the real world, introducing one operator support system more in the control room often involves unforeseen problems. First of all, human operators may feel threatened by such a system if the operators have not been an active part of the design phase. They do not know why and how the system provides support and specifically, why such a system should be able to perform better than themselves. Furthermore, operator support systems like the experience repository presented in this paper requires that the operators share their individual knowledge with a computer system. In general, humans are unwilling to share such knowledge as they are afraid of losing their status to other humans, or as a consequence of the transferring such knowledge to the operator support system. Nevertheless, the attitude in Hydro Polymers AB among the operators has mainly been positive. They have got a complete introduction to the experience repository by one of the process engineers. Moreover, a significant aspect to the successful acceptance among the operators is the participation of and the convincing interest from the plant section management.

Task allocation between the computer system and the human operator

The strength of the computer is to handle large amounts of information and knowledge whereas the human strength is to make decisions and to solve problems. In general, the stages of human information processing include sensing, perception, translation from perception to action and finally, movement. Regarding

Rasmussen's three level model of behaviour, the human solves unfamiliar problems like unexpected situations using knowledge-based behaviour (Rasmussen 1983). This type of behaviour is goal-oriented in that, the problem solver defines a goal based on an identification of the situation and thereafter, makes a suggestion for a solution that fits the goal. Human information processing may take place at various levels of behaviour since new information inputs, or the recognition of a known information pattern, may change the level of behaviour during the process.

Since knowledge-based behaviour is a very demanding process, it is desirable to support the operator at this level of behaviour. On the other hand, the operator becomes a better problem solver by practising this type of behaviour. Hence, it is a matter of providing the operator with useful support while maintaining the operator's situation awareness and skills in performing problem solving. The framework for the proposed operator support system utilises the human ability to:

- identify a new situation
- define a goal for the knowledge-based behaviour
- judge among the retrieved cases
- adapt a previous case to the new situation

In contrast, the computer system performs tasks that require a high capacity of repeated activities and specialities which humans are unable to perform. Hence, the proposed operator support system leaves the computer system to:

- refine the situation description
- retrieve somewhat similar cases from the experience repository
- present the retrieved cases in a 3D visualisation
- maintain the structure of the experience repository

Regarding the human operators' tasks, the first two items refer to the process of recognising and describing a new unexpected problem. Operators describe the new problem situation by observing derivations in the process variables in proportion to individual expectations of the state of these variables. Such derivations are denoted symptoms. The set of symptoms performs features of each particular situation. In some cases, however, the operator is not able to recognise the situation as an already experienced situation based on the symptom(s). Hence, the operator defines a goal for further reasoning. Examples of goals are to diagnose the problem or to minimise the effect, which the symptoms have on other parts of the process. The two tasks described above are essential for the operator to be aware of in order to perform the further problem solving process. In contrast, if the operator support system presents an alarm to the operator that indicates an unexpected situation followed by a potential solution, the operator is normally not able to understand the situation and thus, realises the suggested solution without additional consideration. Therefore, it is significant to include the operator in the process of defining the problem situation.

Nevertheless, in plant operation, operators often observe too few symptoms to describe a new unexpected situation such that it can be distinguished from other situations. Moreover, as the operator already has encoded all observed symptoms, the operator support system therefore supports the operator in the process of refining the description of the new problem situation. The framework for the operator support system incorporates simple rules extracted directly from process data to perform the acquisition of additional symptoms. Techniques within the rough set approach (Pawlak 1991) analyse the original process data and generate rule sets that form distinct decision classes regarding the causes. The rules are simple IF-THEN rules and are generally legible such that the operator quickly can get an overview of knowledge represented by the rules. The proposed operator support system is, however, not supposed to diagnose the situation, but only to support the operators in their decision making processes. The proposed framework utilises the IF-THEN rules to categorise the new and incomplete situation into a certain class. Based on the set of previous cases within this class, the operator support system searches for supplementary symptoms either manually by requesting the operator or automatically by searching in the historical process data.

The computer system is also responsible for the next two tasks to be performed. Based on the refined situation description, the operator support system retrieves a set of similar previous cases from the experience repository. This task involves managing large amounts of information and knowledge and hence, is suited for the computer.

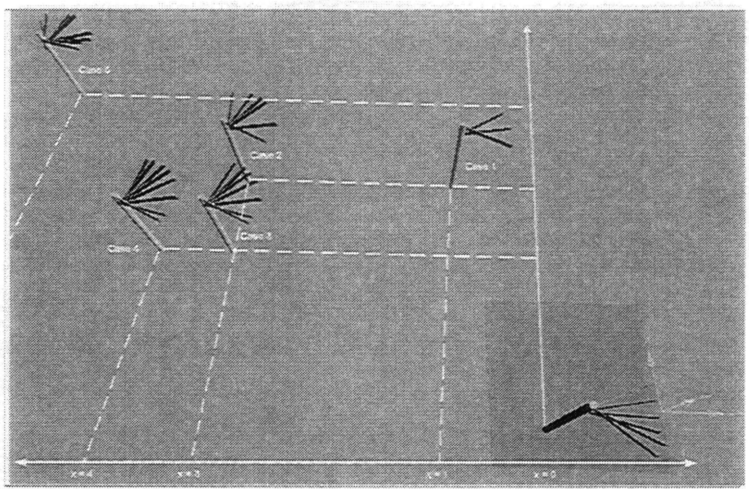


Figure 1: Example of 3D visualisation of retrieved cases

The proposed operator support system furthermore presents the result of the retrieval procedure in a 3D visualisation. This visualisation is meant to be an alternative to the conventional one-dimensional ranked list of retrieved cases. Figure 1 illustrates an example of the 3D visualisation. In a ranked list with a score of the match, the human being very often choose one of the first items as the "right" match (compare with the result of a search on internet). On the other hand, the proposed 3D visualisation presents more dimensions of the retrieval such that the "best" match is not obviously given in the 3D visualisation. However, this visualisation illustrates the relationship among the cases, among the features (such as symptoms, causes and actions) and between a case and its features. The metaphors chosen for this presentation are only meant to illustrate the principles of visualising further relationships.

The human operator is responsible for judging and selecting the most useful match, if any at all, among the retrieved cases. This task requires information processing at the knowledge-based level of behaviour. The operator selects one or more previous cases regarding the context of the new problem situation. In some situations, the number of common features is most relevant whereas, in other situations, the similarity of the common features is the most relevant relationship. Thus, the operator's knowledge about the situation is vital in the process of judging the retrieved cases. Finally, the operator implements a solution in reality that is either based on a similar previous solution, or on redirecting the operator's focus of the new unexpected situation such that he is able to solve the problem without any further help. The operator then encodes the actions, which he has performed, and the outcome of these actions, whether they are successful, to complete the new case.

Support for active information perception from the environment

Human information processing in real-time tasks are extremely situation- and person-dependent (Endsley 1995). The operator actively seeks and picks up information, such as sensor readings and field observations, in the environment that is relevant for the individual's identification of the situation. Based on this situation identification, the human defines goals for the information-processing process. In this particular framework, the support provided by the operator support system aims the operator in two different ways as a retrieved case may represent:

- a previous solution that exactly matches the new unexpected situation
- new aspects of the problem situation which the operator has not been aware of

The first type of support occurs if the specific unexpected situation has been experienced before. In the latter item, the operators get support in terms of an extension of their situation awareness. As the operator support system presents the retrieved cases for the operators, the content of these cases may change the focus of the problem solver and thus, result in directing the situation awareness on other pieces of information that are relevant for the situation and its context. The operator may then be able to solve the unexpected situation on his own without any further help. According to Rasmussen's three level model of behaviour, knowledge-based behaviour becomes an iterative process in that presenting previous cases may address the focus on new information inputs that again may redefine the goals. In addition, the presentation of previous cases extends the operator's individual memory. Hence, the experience repository functions as a common knowledge repository and thus, individual experience becomes available for all the operating staff members.

During discussions with operators and domain experts, they state that members of the operating staff actually reason by retrieving previous cases. Furthermore, operating staff members have suggested situations by themselves for utilising such an experience repository. Some examples are to support information and knowledge transfer:

- in new unexpected situations
- within sections of the process plants where the operators work rarely
- between operators at different skill levels
- to update oneself about occurred unexpected situations after, for example, vacations

To sum up, the operators have been very open-minded and interested in the experience repository as a facility to contain information and knowledge which the operators need to operate the plant, even though they only utilise the proposed framework in certain situations. Also, the focus of supporting operators instead of performing automatic diagnosing makes the operators much more positive as the operator support system only is a tool and hence, does not threaten the operators' status.

CONCLUSION

In the work presented in this paper, a framework for an operator support system is based on case-based reasoning and adapted for the process industry to support unexpected situations. The experience repository using case-based reasoning has been implemented and tested in real environments whereas the 3D visualisation only exists as a proposal to improve the conventional presentation of retrieved cases. However, the purpose of this paper is to allocate the different tasks between the human operator and the computer system. The contribution of the research described in this paper may be summed up as follows:

- The content of the experience repository effectively supports information and knowledge transfer. Hence, the experience repository functions as a common memory for all operating staff members.
- The experience repository provides support for unexpected situations in that, cases contain facts, but also knowledge and intuition, regarding unexpected situations.
- Facing a new unexpected situation, the operator gets support to perform a wider identification of the situation by extending the personal situation awareness during the presentation of previous similar cases.

ACKNOWLEDGMENTS

This work is supported by the Norwegian Research Council and is a part of the Ph.D. research programme, ORD (Operator communication, Control and Distributed computer systems) at Department of Engineering Cybernetics, NTNU. The authors would also like to thank the staff members in all the visited plants who have contributed to this work.

REFERENCES

Endsley, M. (1995). Towards a theory of situation awareness in dynamic systems. *Human Factors. 37*, 32-64.

Kitano, H., & Shimazu, H. (1996). The experience-sharing architecture: A case study in corporate-wide case-based reasoning software quality control. In D. B. Leake (Ed.), *Case-Based Reasoning – Experiences, Lessons, & Future Directions (235-268).* Menlo Park, California: AAAI Press.

Kolodner, J. (Ed.). (1993). *Case-Based Reasoning.* San Francisco: Morgan Kaufmann.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science. 5*, 14-37.

Pawlak, Z. (Ed.). (1991). *Rough Sets – Theoretical Aspects of Reasoning about Data.* The Netherlands: Kluwer Academic Publishers.

Rasmussen, J. (1983). Skills, rules, and knowledge; Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics. 13*, 257-266.

Skourup, C., and Alty, J. L. (1998). A case-based reasoning approach to fault diagnosis in control rooms. Paper presented at the *7th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems.* Kyoto, Japan, Sept.

RÉSUMÉ

ADAPTATION DES EXPÉRIENCES OPÉRATIONNELLES POUR LES SITUATIONS IMPRÉVUES

Cette communication présente les grandes lignes d'un répertoire d'expériences en vue de gérer des observations faites par les opérateurs sur le processus, tout autant que des intuitions et des connaissances sur des situations imprévues. Dans la réalité, l'introduction d'un système de soutien à l'opérateur de plus en salle de contrôle conduit souvent à des problèmes inattendus. Le cadre qui est proposé exploite les qualités respectives de l'ordinateur et de l'opérateur humain pour partager les tâches entre les deux agents. Dans la mesure où l'activité fondée sur les connaissances déclaratives est très exigeante, il est souhaitable de soutenir l'opérateur à ce niveau. Chez l'homme, le traitement d'information dans des tâches exécutées en temps réel est très dépendant de la situation et de la personne.

MOTS CLÉS : Assistance à l'opérateur, Répartition de tâches, Fonctionnement d'une installation, Assistance visuelle à la décision.

Session 3: Cognitive cooperation
Coopération cognitive

Human/Human and Human/Systems interactions in highly automated cockpits. Instructors' interventions during full size simulator sessions in qualification training.

Marielle Plat & Janine Rogalski
Laboratoire Cognition & Activités Finalisées
ESA 7021 CNRS-Université Paris 8
2 rue de la liberté,
F-93526 Saint-Denis Cedex2.
marielle.plat@univ-paris8.fr rogalskij@univ-Paris8.fr

ABSTRACT

The study concerns the way instructors manage interactions between pilots and interactions between crew and systems, with regards to difficulties pilots encounter with automated systems (AS). The empirical study took place during simulation sessions on Full Size Simulator (FFS), at the end of training soon experienced pilots for type qualification, for highly automated cockpits. Instructors' interventions are considered from two points of view. Firstly, they may reveal expected difficulties concerning human/human/systems interactions. Secondly, they are indicators of what is dominating didactical interventions on pilots' activity during full size simulations. The results showed that, in fact, more interventions concerned AS than other systems. Globally, they were expressed during crews' action, for correction or scaffolding. However, during the approach phase, instructors intervened more often "out of action" (before or after) when concerned by AS. Most of the interventions were reinforcing crews' action or giving an operative status to previously acquired knowledge on AS. There was an evolution with trainees experience and efficiency in FFS sessions: negative interventions were strongly decreasing. In spite of a high variability between crews there were invariants: human/ human co-operation appeared to be considered as "Cinderella", and instructors were mainly focused on actual crew performance; they gave few hints to control how AS were functioning and to identify cues for this purpose.

INTRODUCTION

Before presenting the empirical study, we will recall the main characteristics of aircraft piloting as process control, with regards to human reliability related to co-operation within crew, and to interactions with systems and particularly with automated systems. The main lines of the study are then sketched. The main results concern importance, nature and evolution of instructors' interventions through FFS sessions.

Piloting with highly automated cockpits is highly demanding both in Human-Human and Human/Systems interactions. These interactions were identified as key points for reliability. Aircraft piloting can be characterised as managing a dynamic environment presenting several features of complexity: very fast "tempo"; high level of risk (particularly during Take-off and Approach/landing); indirectness of information taking and of control—mainly due to the role of automated systems; rich and highly complex scope of supervision, implying spatio-temporal variables. This leads to a high level of complexity of aircraft piloting, whatever the categorisation of process control situations (Woods, 1988; Hoc, 1993).

Complexity is still increased, from the one hand, with interdependency of pilots' activity in the crew, and, on the other hand, with mediation of crew's actions through automated systems (AS from now on). AS may be considered as system of agents involved in cockpit interactions, so that two types of interactions are involved: within crew and between crew and AS. Both types of interactions became central in aviation studies in the last years. The needs for supporting and training human co-operation were underlined in research papers from the nineties, (Green, 1990), while studies are now developing about how to ensure team situation awareness through system design and/or training (Endsley, 1995; Schretha et al., 1995).

In fact, AS act as agents "at a high level of autonomy and authority [...] they are capable of modulating or overriding user input authority" (Sarter & Woods, 1995b, p.1); they operate as active filters between crew and aircraft, with a low level of readability, far from the transparency Hutchins claimed as necessary for efficient interactions (Hutchins, 1990; Hutchins & Klausen, 1991). Those systems imply high crew interdependency but they could also create individual behaviour by their private access possibilities. Wiener (1985), Amalberti & Valot (1987), emphasised that wrong representation of the situation is related to a too limited understanding of AS with their particular logic. Wiener (1989), James & al. (1991) sustained that missing feed-back is contributing to this misunderstanding. Now, crew situation awareness requires a high level of common awareness of AS dynamic status, to avoid mode errors (Sarter & Woods, 1994; 1995a), to face "automation surprises" (Sarter & Woods, 1995b), and other unexpected elements requiring a "conceptualisation power" from crew on AS (Woods et al., 1990). A recent experimental study, on how competent crews face unexpected AS dysfunctioning, show that pilots cannot help trying to understand the situation, even when procedures require to come back to manual control in such cases (Plat et al., 1998 ; Plat & Amalberti, to appear).

High requirements are expressed concerning on-line crew co-operation through "shared control and check". In spite of this, limitations are observed in human/human interaction due to common schemes of communication (Rogalski, 1996). These schemes act as obstacles, and in crew/AS interactions, as concerning AS status, behaviour and intentions—what has been programmed in AS before the flight, as well as entered by pilots as instructions during the flight. Pilots' knowledge, acquired through training, seems to function as "inert knowledge" (Sarter & Woods, 1994) when confronted to strong real time constraints in situation assessment.

The concern of the study presented below is to analyse the role played by direct instructors' interventions on these sensible points, when they manage high fidelity simulations at the end of qualification training, focusing on the properties of systems and procedures specific to a highly automated cockpit type. One key issue if the following: do instructors' interventions allow crews to contextualise knowledge acquired during theoretical training about automated systems. Such a contextualisation should be indicated by interventions concerning understanding automated systems functioning, and should contribute to transform “inert knowledge” into operational one.

PRESENTATION OF THE STUDY

The data of this study were transcriptions from actions and communication within “ instructor/crew ” triads (6 different crews), on a full flight simulator at the end of a A320 certification. Two types of situations were recorded: an important incident in a critical phase (engine fire during Take-off) and a normal instrumented approach, with disruption at landing (go-around). A first series of results concerned co-operation activity within the crew, and instructors' interventions during incidental take-off (Rogalski et al, 1994; Rogalski, 1996). Studying transitions between the main task (Take-off) and the critical incident (engine fire) showed that explicit shared control was first achieved on announcements (commands for action) and information about action execution, and was less reliable when concerning information about the situation. Better co-operation followed better (procedural) crew performance. There was no specific concern about crew/AS interactions, but pilots' activity showed that giving a new "law of flight" to AS was the most difficult transition. Concerning instructors' interventions, they appeared to be first centred on procedural performance, then on co-operation.

In the present study, data were analysed from the point of view of instructors' interventions concerning crew/systems interactions, and within crew interaction. Incidental Take-off and Approach phases were compared, with an emphasis on the relative place of As with regards to other systems. Evolution through simulation exercises was analysed through a comparison of three sessions (around 30 minutes each) at the end of the qualification training (sessions 3.1; 3.2 and Session 6 preceding the qualification test).

METHODOLOGY FOR DATA ANALYSIS

Instructors' interventions were analysed from the following points of view :

- nature of the systems involved in current crew interaction : automated systems/other systems
- moment of the intervention with regards to crew actions on the systems: before action, during action, after action; the topic of instructor's intervention was used as a critical cue for identifying this moment.
- value of the intervention: positive—when instructors were reinforcing crew actions or teaching new knowledge—, or negative—when they were correcting crew actions.
- nature of the intervention: interventions were categorised as “operative”, “cognitive” (concerning crew / systems interactions) or “co-operative” (about co-operation in human / human interaction): Operative interventions help crews to co-operate with the system: the main aim is improving performance: Cognitive interventions are oriented towards improving crews' understanding of AS; co-operation intervention are oriented towards human/human interactions. Interventions might also concern the simulation session itself.

  Operative

  - preventive interventions concerning automated systems (action that crew must certainly not do)
  - mnemonic help for using automated systems
  - help to intervene on automated systems while acting
  - help about requirements in order of actions (prerequisite, aso...)

- help to use information given by AS
- correction of action through or on AS

Cognitive

- help to confront actual AS state and previous crew's action
- help to adapt AS use to aircraft state as a mobile
- comment on how to use automated systems (long term anticipation or general remarks)

Co-operation

- drawing crew's attention towards communicating informations and actions linked to AS with the goal of building a shared situation awareness
- recalling requirements for crew's co-ordination when interacting with AS.

Instructors' interventions were also analysed depending on two situation variables:

- flight phases
  - incidental take off
  - approach
- simulation session
  - third sessions of full size simulation: 3.1 and 3.2, where Captain and First Officer were respectively Pilot Flying
  - sixth session: this session was the last one before qualification assessment (real flight).

MAIN RESULTS

Globally, there were 344 for 18 sessions (3 for each of the 6 crews), each session being completed in around 30 minutes. The mean density of instructors' interventions can be considered as relatively high. However, as further results will show, the variability was high with regards to several variables. Table 1 presents the effect of three of these variables on the number of instructors' interventions.

Table 1 : Number of instructors' interventions depending on the nature of the involved systems (automated *vs* others), and on the flight phase (incidental take off *vs* approach)

Systems involved	**Automated systems**		**Non automated systems**		
PHASES	TAKE OFF	APPROACH	TAKE OFF	APPROACH	
Moment					Global
Before action	10	52	8	29	99
After action	4	24	5	16	49
During action	30	67	25	74	196
Global	44	143	38	119	344

Nature of the Systems Involved in Instructors' Interventions

More interventions involved automated systems than others components (54,4% *vs* 45,6%). This was similar for the two flight phases.

Moment of Instructors' Interventions with Regards to Crews' Action

Instructors intervened dominantly during action (57% of the global number of interventions). The moment of intervention depended on the nature of the systems: Instructors intervene more often before or after action when automated systems were involved than when other systems were concerned. Contrasting "during action" against "out of action" for the two types of systems gave Khi2 (df=1, 344) = 4,37 $p<.05$. This effect was mainly due to the existence of more interventions " before action " for automated systems. There were few differences concerning interventions " after action " (which were always less than 15%).

The distribution of interventions before, during, and after action was slightly depending on the flight phase: they were more often expressed during action in the incidental take Off phase than in Approach phase (67% *vs* 54%).

There was an interaction between phases and nature of the systems involved: the distribution of interventions with regards to action was similar for Take Off, whatever the systems; at the contrary, in the Approach phase, instructors intervened more often "out of action" on automated systems (53%), and less often for other systems (38%).

Value of Instructors' Interventions

Most of the instructors' interventions (67%) were positive: they were reinforcing crews' actions of giving an operative status to the knowledge acquired on automated systems during former training. The difference between automated systems and other systems was limited (70% of positive interventions for AS *vs* 65% for other systems), but present for all sessions, and for the two flight phases. Table 2 shows their evolution through training on full size simulator.

Table 2 : Number of positive and negative instructors' interventions depending on the sessions.

value of interventions **Session**	**positive**	**negative**	**Total**
Session 3.1	65	50	**115**
Session 3.2	84	36	**120**
Session 6	83	26	**109**
Global	**232**	**112**	**344**

The density of instructor's interventions did not vary largely during the three sessions (a little less in session 6). The number of negative interventions was decreasing (44 % in session 3.1, 30% in session 3.2, 24% in session 6). This trend was similar for automated systems and non automated systems.

These global results conceal dramatic differences between crews, in terms of total number of interventions, and of evolution of positive and negative interventions.

Inter-crew variability

Table 3 : Evolution through training sessions for the six crews

	Positive interventions			**Negative interventions**			**Global**
sessions	3.1	3.2	**6**	3.1	3.2	**6**	+ / -
Crew 3	5	3	**1**	1	0	**1**	**9 / 2**
Crew 2	13	26	**14**	2	3	**3**	**53 / 8**
Crew 6	10	9	**0**	12	7	**0**	**19 / 19**
Crew 1	10	10	**19**	10	2	**1**	**39 / 13**
Crew 4	10	15	**15**	12	6	**8**	**40 / 26**
Crew 5	17	21	**34**	13	18	**13**	**72 / 44**

These data showed strong individual differences between crews:

- For two crews (C3&C6) there were almost no more intervention during the last training session.
- For two other crews (C2&C1), there was almost no more negative interventions, but a lot of positive ones.
- For the two other crews, the instructors intervened a lot even in the last session, including for correcting inappropriate actions (C5&C4). In fact, for the last crew (C4) another training session was used before qualification.

Nature of Instructors' Interventions Concerning AS

The distribution of interventions concerning automated systems was quite similar for sessions 3.1 and 3.2. Results presented in Table 4. contrasted these sessions to the last simulation session 6.

Table 4 : Nature of instructors' instructions concerning automated systems depending on the session.

Sessions Nature of interventions	Sessions 3.1 &3.2	Session 6	Global
Operative	**91**	**30**	**121**
Cognitive	**35**	**14**	**49**
Co-operation oriented	**7**	**4**	**11**
Global	**133**	**48**	**181**

- Interventions involving crew coöperation were quite limited (around 6 percent of all interventions).
- Operative interventions were more present than cognitive ones (around 2/3 versus 1/3). Concerning AS, instructors' interventions were mainly centred on operative acquisition during qualification training.
- Evolution through FFS sessions of simulation was important with regards to the density of instructors' interventions on AS
- there was a quite limited evolution in distribution between operative, cognitive and co-operative interventions: the part of operative interventions decreased from 68,4% to 62%, while there was a slight increase in the part of cognitive interventions (from 26,3% to 29,2%) and of co-operation oriented interventions (5,3% to 8,3%).

DISCUSSION

Analysing instructors' intervention during full size simulator training, was aiming at identifying the main focus of their didactical concerns. Results showed that, in fact, more interventions concerned AS than other systems. Globally, they were expressed during crews' action, for correction or scaffolding. However, during the approach phase, instructors intervened more often "out of action" (before or after) when concerned by AS. Most of the interventions were reinforcing crews' action or giving an operative status to previously acquired knowledge on AS. There was an evolution with trainees experience and efficiency in FFS sessions: negative interventions were strongly decreasing, with high variability between crews. However, instructors' interventions kept invariant characteristics: efficiency of crew' action was more important than quality of understanding. Moreover, it seemed as if crew co-operation were a secondary task, with respects to task performing. Even for AS, human/ human co-operation appeared to be considered as "Cinderella" during on-line interventions.

Results lead to the hypothesis that during FSS sessions, instructors were mainly focused on actual crew performance; they gave few hints to control how AS were functioning and to identify cues for this purpose. Such a focus on performance was in fact observed in other situations involving full size simulations, aiming at producing by trainees an integration of "formal" knowledge at an operative level. Inter-crews variability might partly explain that only a tendency was observed in a decrease of direct operative interventions: Operative interventions should remain necessary for crews performing not quite well when interacting with automated systems. It might also be possible that we observed a didactical effect. Instructors might want to be quite sure about the level of pilots/AS interactions reached by the crew they had in charge: in fact, the last session of full size simulation was just preceding the qualification test. (Actually, one crew had to perform a second time this last session).

It seems that contextualising knowledge about how to interact with automated systems was supposed to be acquired without explicit interventions during simulation and/or through practising. Another hypothesis is an underestimation of knowledge required for dealing with automated systems. Whatever the reasons of instructors' didactical practice, the consequence could be the same: pilots must construct this knowledge themselves. In fact, personal constructed knowledge lets pilots in a state of high cognitive uncertainty to cope with "automation surprises" (Sarter & Woods, 1997), or with dysfunctioning of automated systems (Plat & Amalberti, under press; Plat & Rogalski, 1999). The use of personal knowledge may lead to analogical reasoning—as observed by Plat and al— with a risk of inappropriate actions on a faulty automated system.

REFERENCES

Amalberti, R. (1996). *La conduite des systèmes à risque*. Paris : PUF.

Amalberti, R., Valot C. (1987). Erreurs humaines dans certains accidents aériens et perspectives d'"aides intelligentes". *Rev. Med. Aéro. & Spatiale*, (26) 104, 319-320, 1987.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37 (1), 32-64.

Hoc, J.-M. (1993). Some dimensions of a cognitive typology of process control situations. *Ergonomics*, 36 (11), 1445-1455.

Green, R. (1990). Human errors on the flight deck. In D. E. Broadbent, J. Reason & A. Baddeley (Eds.), *Human factors in hazardous situations* (pp. 503-512). Phil. Trans. R. Soc. London. B327.

James M., Mc Clumpha A., Green R., Wilson P., Belayn A., (1991). Human Factors for pilots. *XIX th WEAA*. Nice 1991, 91-106.

Hutchins, E. (1990). The technology of team navigation. In J. Galegher & C. Kraut (Eds.) *Intellectual team work* (pp. 191-220). Hillsdale NJ: LEA.

Hutchins, E. & Klausen, T. (1991). Distributed cognition in an airline cockpit. In Y. Engeström & U. Middelton (Eds.), *Cognition and Communication at work*. Cambridge: Cambridge University Press.

Plat, M. & Amalberti, R. (on press). Crew training to automation surprises, In N. Sarter & R. Amalberti (Eds.), *Cognitive Engineering in the Aviation Domain*, Hillsdale NJ: LEA.

Plat, M. & Rogalski, J. (1999). Compréhension de dysfonctionnements d'automatismes et modes de coopération dans le cockpit. In H. Benchekroun & A. Weill-Fassina (Eds.), *Approches d'analyse du travail collectif.* Toulouse: Octarès.

Plat, M., Rogalski, J. & Amalberti, R. (1998). Pilot-automation interaction and cooperation in highly automated cockpits. In G. Boy, C. Graeber & J.-M. Robert (Ed.), *HCI Aero'98* (pp. 229-234). Montreal, Canada: Editions de l'Ecole Polytechnique de Montréal.

Rogalski, J. (1996). Co-operation processes in dynamic environment management: Evolution through training experienced pilots in flying a highly automated aircraft. *Acta Psychologica*, 91 , 273-295.

Rogalski, J., Samurçay, R. & Amalberti, R. (1994). *La coopération et la coordination dans les cockpits automatisés.* Final Report DGAC / CNRS-Université Paris8-CERMA.

Rouse, W. B., Cannon-Bowers, J. A. & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE transactions on systems, man and cybernetics*, 22 (6), 1296-1308.

Sarter, N. B. & Woods, D. D. (1992). Mode error in supervisory control of automated systems. *Human Factors Society 36th Annual Meeting* . Atlanta, GA.

Sarter, N. B. & Woods, D. D. (1995a). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, 37 (1), 5-19.

Sarter, N. B. & Woods, D. D. (1995b). *"Strong, silent, and 'out of the loop' ": Properties of advanced (cockpit) automation and their impact on human-automation interaction.* CSEL.

Sarter, N. B. & Woods, D.D. (1994) Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the Flight Management and Guidance System. *International Journal of Aviation Psychology*, 4(1), 1-28.

Sarter, N. B. & Woods, D. D. (1997). Team play with a powerful and independent agent: operational experiences and automation surprises on the Airbus A-320. *Human Factors*, 39(4), 553-569.

Schretha, L. B., Prince, C., Baker, D. P. & Salas, E. (1995). Understanding situation awareness: Concepts, methods and training. In W. B. Rouse (Ed.) *Human / Technology Interaction in Complex Systems* (pp. 45-83). Greenwich CT: JAI.

Wiener, E. (1985) Beyond the sterile cockpit, *Human factors*, 27 (1), 75-90.

Wiener, E. (1989). *Human factors of advanced technology ("glass-cockpit").* Transport Aircraft, NASA Contractor Report 177528.

Woods, D. D. (1988). Coping with complexity: The psychology of human behavior in complex systems. In L. P. Goldstein, H. B. Andersen & S. E. Olsen (Eds.), *Mental Models, Tasks and Errors,* (pp. 128-148). London: Taylor & Francis.

Woods, D. D., Roth, E. M. & Bennett, K. B. (1990). Explorations in joint human-machine cognitive systems. In S. P. Robertson, W. Zachary & J. B. Black (Eds.), *Cognition, Computing, and Cooperation.* (pp. 123-158.). Norwood NJ: Ablex Publishing Corporation.

RÉSUMÉ

LES INTERACTIONS HOMME-HOMME ET HOMME-SYSTÈME DANS LES HABITACLES TRÈS AUTOMATISÉS. L'INTERVENTION DE L'INSTRUCTEUR DANS LES SESSIONS DE SIMULATEUR PLEINE ÉCHELLE AU COURS DE LA FORMATION À LA QUALIFICATION.

Cette étude porte sur l'analyse des interventions des instructeurs en situation de formation et leur gestion des interactions entre pilotes, et entre pilotes et systèmes. Nous avons pris en compte tout particulièrement les difficultés que les pilotes rencontrent avec les systèmes automatisés. Les données empiriques ont été produites à partir de sessions en simulateur pleine échelle durant une qualification de type sur avion *glass-cockpit*. Les interventions des instructeurs ont été considérées de deux manières. Tout d'abord, comme révélant des difficultés concernant les interactions Pilote/pilote/systèmes. Par la suite, comme indicateur du contenu des interventions didactiques de l'instructeur, en liaison avec l'activité des pilotes durant les sessions sur simulateur pleine échelle. Les résultats montrent qu'il existe plus d'interventions qui concernent les interactions avec les systèmes automatisés qu'avec les autres. Les interventions des instructeurs sont effectuées généralement durant l'action des pilotes, pour corriger ou apporter une aide à l'action. Néanmoins, au cours de la phase d'approche on remarque que les interventions s'effectuent le plus souvent en dehors de l'action de l'équipage (avant ou après) lorsqu'elles concernent les interactions avec les systèmes automatisés. La plupart des interventions viennent renforcer des actions de l'équipage ou opérationnaliser des connaissances théoriques (supposées acquises) sur les systèmes automatisés. On a pu mettre en évidence une évolution des interventions avec l'expérience et de l'habileté des pilotes durant les sessions d'entraînement : les interventions négatives diminuent significativement. Malgré une grande variabilité inter-équipages il existe des invariants : la coopération entre pilotes semble être traitée en "Cendrillon" ; les instructeurs sont focalisés sur la performance immédiate des équipages. Enfin, les instructeurs ne donnent que peu d'indications permettant le contrôle du fonctionnement des systèmes automatisés afin que l'équipage sache repérer ce que fait le système et son adéquation aux intentions.

ASSESSMENT OF A METHOD TO STUDY COGNITIVE COOPERATION IN FIGHTER AIRCRAFT PILOTING

Anthony LOISELET and Jean-Michel HOC
Anthony.Loiselet@univ-valenciennes.fr Jean-Michel.Hoc@univ-valenciennes.fr
CNRS - UVHC
LAMIH - PERCOTEC
Le Mont Houy – B.P. 311
F-59304 VALENCIENNES CEDEX

ABSTRACT

This text presents the main results of a preliminary study of human-human cooperation (and its assistance) in a two-seater fighter aircraft. From behavioural and (simultaneous and self-confrontation) verbal data obtained during the observation of a test pilot team on a part-task simulator, an individual protocol analysis has been done. This analysis has allowed us to assess the relevance of a coding scheme of cooperative activity stemming from the theoretical framework of Hoc (1998). It has also allowed us to adjust the scheme and to highlight the main features of cooperation in fighter aircraft. The latter will be used to propose some support the study and evaluation of which will be the next step of the research project.

Keywords: Cognitive Cooperation, CSCW, Individual Protocol Analysis, Coding Scheme, Fighter Aircraft Piloting.

INTRODUCTION

Within the context of a scientific collaboration between Dassault-Aviation and the CNRS, on human-human cooperation in fighter aircraft, a first study of cognitive cooperation between Pilot (P) and Weapon System Officer (WSO) has been conducted in a simulator of two-seater fighter aircraft. The aim of this first study was mainly (a) to assess the relevance of an individual protocol analysis method, and (b) to elicit the main features of human-human cooperation in the cockpit. The next step, going on with theoretical and methodological issues, will be to define and study CSCW assistance for human-human cooperation.

Our approach to cooperation is more operational than structural. We are trying to describe cognitive activities embedded in cooperation rather than structural relationships between agents. This kind of approach is particularly relevant to the study of cooperation in teamwork as opposed to groupwork. Teams are composed of a small number of deeply task-oriented agents (Salas, Prince, Baker, & Shrestha, 1995) whereas groups can be larger and more weakly concerned by a common task. Thus, group processes like group coherence, trust, etc. (Kræmer & Pinsonneault, 1990), which are not prominent in teamwork, are not studied. We are focusing on data relevant to guide the design of assistance to human cooperative activities.

THEORETICAL FRAMEWORK

Cognitive Cooperation: a Definition

Cooperation as a cognitive activity can develop at two conditions (Hoc, 1998):

(i) Each agent strives towards goals and can interfere with others (on goals, resources, procedures, etc.).
(ii) Each agent tries to detect and manage such interference to make the individual and collective activities easier.

The notion of interference is borrowed from studies on planning (Castelfranchi, 1998; Hoc, 1988). It is opposed to independence, but can take diverse forms, like:

- Precondition relationships: an agent's goal is considered as a precondition for another agent's goal.
- Interaction relationships: two procedures distributed among two agents jeopardise each other if each one is not designed considering the other.

At first glance, the effect of interference in collective activity could be considered negatively. However, a deeper analysis reveals that interference can result in positive outcomes, for example:

- Improvement of the team adaptive power (e.g., integrating different points of view).
- Improvement of performance (e.g., criticising the partner after mutual control).

A cognitive architecture

Our approach to cooperative activities began by exploring three abstraction levels before studying the relationships between them. One of the main results of this observational study was the identification of some of these relationships. Apart a kind of metalevel (elaborating compatible representations, elaborating a model of oneself and of the others agents), two levels are stressed here:

Cooperation in Action

This level of activity enables the agents to manage interference locally, integrating cooperation into the course of action, sometimes very quickly. Interference creation (ITF-CR) (e.g., following mutual control), detection (ITF-DT), and resolution (ITF-RS) can develop following almost reactive strategies according to circumstances. However, identifying the other agents' goals (GOAL-ID) introduces possible anticipation in managing interference. At this level, such identification is supposed to rely on domain knowledge only. However, at a higher level, the availability of a model of the other agent can improve this identification, but is not always necessary.

Cooperation in Planning

This abstraction level is needed, both when local management of interference is not efficient and when this kind of failure can be anticipated (e.g., during mission preparation). At this planning level, the elaboration or maintenance of a Common Frame Of Reference (COFOR) is required. This concept is close to those of mutual situation awareness (Salas *et al.*, 1995) and common ground (Clark, 1996). It can be seen as an extension at the team level of the notion of problem space (Newell & Simon, 1972) or current representation (Hoc & Amalberti, 1995). It plays a major role in communication understanding by providing a shared context constraining the interpretation of messages (focal information). It orients towards more accurate representations of problems within the context of a collective activity. Three kinds of activity can be defined:

(i) Elaborating a common goal or plan (CO-GOAL)

These entities can contribute to avoiding interference by providing a common organisational structure for action. Overall coordination of actions can be processed at this level instead of local coordination that may resolve interference only temporarily.

(ii) Elaborating role allocation (ROLE)

Adaptation to changing situation very often introduces the need for changing a prior distribution of roles among agents into a dynamic role allocation.

(iii) Elaborating other COFOR components (COFOR)

For the purpose of this preliminary study, we isolated only common goal, common plan, and role allocation from other COFOR's components like environment state (or change), resources, operator activities, etc.

METHOD

Data Collection

The data comprised results from a single observation of a test pilot team (P and WSO) performing the central part of a military mission (a ground target attack) in a *Mirage 2000*-type generic part-task simulator. The simulator and the test pilot team were provided by Dassault-Aviation and the mission was designed in collaboration with Dassault-Aviation's test engineers. The observation was divided into five steps:

1. *Briefing.* A test engineer briefed the team about the mission it had to complete and we asked the operators to explain the way they planned to perform the task together. In this stage, we recorded the operators' verbal reports.
2. *Mission.* The team completed the mission. We used two video cameras to record three interfaces presented in the cockpit (P's head-up display and WSO's two lateral displays) and an audio-recording system to record the communication in the crew.
3. *Self-confrontation.* Just after the mission, we performed a self confrontation of the team (presenting the recordings of the previous step) and recorded the verbal reports of the operators commenting their own activity.
4. *Interviews on cooperation.* Two hours after, the operators were interviewed together about some details of their cooperative activity. Once more, verbal reports were recorded.
5. *Interviews on the work domain.* Finally, two weeks after the observation, we performed an interview of the pilot to get some missing information to help us to infer some covert activities.

Individual Protocol Encoding

This method of analysis involved inferring the operators' cognitive activity from different behavioural (verbal and non-verbal) data and coding it using a predicate-argument formalism. A predicate codes a process and its arguments, the representation processed and some specification of the process. The inferences are based on a cognitive architecture, on the work domain knowledge, on the activity context, and on the operators' verbal reports during the mission and self-confrontation. We mainly used two kinds of cognitive architecture:

- DSM (Dynamic Situation Management) to code individual diagnosis and decision-making activities (Hoc & Amalberti, 1995);
- The framework presented above to code cooperative activities or the cooperative function of individual activities (a cooperative activity is obviously performed by an individual).

Here we will stress the coding scheme used to code cooperative activities. The scheme contains one predicate for each of the activities presented above. Table 1 presents the arguments common to all the cooperative activity predicates.

ARGUMENT	DEFINITION
Cooperative activity specification	specifies the type of processing (e.g., the type of interference created)
Means type	indicates the means by which the cooperative activity is performed (e.g., by a verbal communication)
Means value	specifies the means (e.g., the type of verbal communication used: a question, an acknowledgement, etc.)
Data	codes the data the activity processes
Results	codes the consequences of the cooperative activity on the individual, on the other agent, on the team, or on the environment
Underlying activities	specifies the individual activities underlying the cooperative activity
Agent	specifies which member of the team achieves the cooperative activity
Condition	codes the conditions without which the cooperation cannot appear
Goal type	codes whether the goal of an activity is to make oneself, partner's or team's activity easier
Goal value	specifies the goal of the cooperative activity

Table 1: Arguments of cooperative activity predicates

RESULTS AND DISCUSSION

Assessment of the Coding Scheme Relevance

Types of Interference

The data show four types of interference that are managed in different ways and for different purposes (depending on the nature of the work situation):

(i) A precondition interference is mainly identified by verbal orders that are given to the partner who must respond to it by a verbal acknowledgement or by doing the action. From that point of view it is a positive interference. An operator detects a negative precondition interference when the conditions for an action are not met by the partner.

WSO simultaneous verbal reports	Cockpit activity explanations	Coding of WSO cooperative activity *Predicate (activity specification)*
	WSO identifies that the aircraft is detected by a dangerous enemy threat	
"ok, break left"	WSO decides that a break manoeuvre is needed and prompts P to do it.	ITF-CR (precondition)

(ii) Interaction interference is negative and must be detected and resolved. When created, it is not purposeful but an unanticipated consequence of an intended action.

WSO simultaneous verbal reports	Cockpit activity explanations	Coding of WSO cooperative activity
	After a break (before the way-point 31), the aircraft must return on the flight plan. In order to optimise the path of the aircraft the WSO decides to directly rejoin the way-point 32 (instead of going back to the way-point 31).	
	There is an interaction interference between the managing of the path to follow by the WSO and P's activity (through the configuration of the steering indicators).	ITF-DT (interaction)
"Ok, I select you the way-point 32 right now"	WSO announces his action. This statement contributes to the anticipated solving of the interference.	ITF-RS (interaction)
	WSO selects the way-point 32 as destination way-point	ITF-CR (interaction)

(iii) Mutual control interference is only created with the goal of improving the partner's performance (positive interference).

WSO simultaneous verbal reports	Cockpit activity explanations	Coding of WSO cooperative activity
	P stops the break and flies straight ahead.	
"Continue like that"	WSO evaluates the situation and concludes that the pilot's action is suitable.	ITF-CR (mutual control)

(iv) Redundancy interference appears when the two operators' activities focus on the same task simultaneously (not necessarily consciously). This type of interference increases the cockpit's adaptivity and reliability but it can degenerate into negative interference if not adequately managed.

WSO individual activity	P individual activity	Cooperative activity coding
	P is doing a break to escape from an enemy threat (in response to a WSO's order).	
WSO is monitoring the evolution of the threat to determine the postcondition of the break.	P is monitoring the direction of the aircraft in order to determine the postcondition of the break.	ITF-CR (redundancy)
The first identifying one of the precondition decides to stop the manoeuvre.		

Despite the fact that some types of interference may be most of the time either positive or negative, it is the way they are managed by the cockpit's activity that finally makes them positive or negative (and not their nature).

Cooperation in planning

The second level of cooperation mainly concerns the management of overall representational structures. However, this management takes two forms in the studied protocol (illustrated by the COFOR predicate):

(i) The first one concerns the elaboration of a new representation structure or a large modification of a pre-existing one. This activity creates a specific episode that interrupts (or overlaps) the course of activity.

WSO simultaneous verbal reports	P simultaneous verbal reports	Cooperative activity coding *Predicate (activity specification, means type, means value, agent)*
	"The LDP (*instrument allowing the guidance of the bomb*), what is it aiming at?"	COFOR (elaboration, verbal, query, P)
"The PDL ... it's the mode, then it's all right. It's inertial pre-aiming, we're on 33"		COFOR (elaboration, verbal, answer, WSO)
	"It's... it's aiming at the target?... No? Not yet"	COFOR (elaboration, verbal, query, P)
The discussion is continuing		

(ii) The second one concerns the continuous updating of pre-existing structures. This activity is integrated into the course of activity.

WSO simultaneous verbal reports	P simultaneous verbal reports	Cooperative activity coding *Predicate (activity specification, means type, means value, agent)*
"xx miles"		COFOR (maintenance, verbal, assessment, WSO)
"We're soon going to enter the firing range"		COFOR (maintenance, verbal, operational translation, WSO)
	"Ok"	COFOR (maintenance, verbal, acknowledgement, P)

This distinction is particularly relevant for COFOR (the main activity found in the protocol at this abstract level). It is relevant because the features of these two forms of activities are different, their implications on the individual activity of the operators are different and their relationships with the other cooperative activities are also different.

Making the definitions more precise

During the application of the scheme to the protocol data, it appeared that the different cooperative activities were not defined exclusively. This imprecision produces difficulties in coding the protocol.

These difficulties mainly stress two types of relationship between the different activities of the scheme:

- Sometimes, cooperation in action (local and concrete level) can also result in modification of more abstract representational structures (global level). For example, activities of creation, detection or resolution of interference can update COFOR.
- Conversely, high-level activities determine low-level ones. For example, updating COFOR will always create interference on the partner's activity and can also resolve the interference.

Role of communication in cooperation

Three points seem interesting to note as regards our results:

- We are faced by similar problems to those met within the framework of the study of communication. Especially any communication constrains the space of possible interpretations but does not provide a precise and single meaning. The same observation could be made about cooperative activities that rest on a verbal means.
- 78% of cooperative activities are conducted verbally (figure 1). However, this percentage could be overestimated, since it is more difficult to infer a cooperative activity without utterance.
- Our results show that a study of cooperation soley based on verbal communications is not sufficient. Distributions of cooperative activities are different if one considers all these activities or only the part that uses a verbal means (figure 1). Only taking into account activities which are conducted verbally results in a selective filtering of activities over the different classes of the scheme. This filtering is particularly true for interference detection and the creation of mutual control, interaction and redundancy interference. Note that cooperation in planning is always conducted verbally.

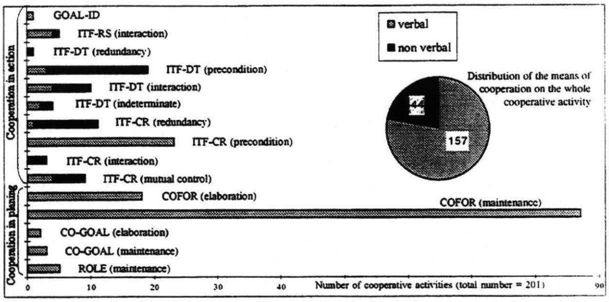


Figure 1: Distribution of the mean of cooperation in the different cooperative activities

Results on Cooperation in the Cockpit

The assessment of the coding scheme by its use, as mainly done in the previous section, is not sufficient. In this section, we will give some examples of results that can be expected from the application of this scheme to data on cooperation.

(i) Our analysis points out operator's needs for information about the team's activity. It is the main topic of communication (53%) and the management of mutual control and redundancy interference matters a great deal to the operators in the fulfilment of the mission.

(ii) The method allows an internal evaluation of cooperation by providing information about its structure in the cockpit (as a whole) and in the activity of each members of the team (figure 2):
- There is a fair distribution between cooperation in action and in planning in the cockpit.
- WSO is more cooperative than P (WSO performs more cooperative activities, is more oriented towards P's activity and its easiness, communicates more).
- COFOR management (apart from ROLE and CO-GOAL) has a prominent role in cooperative activity.

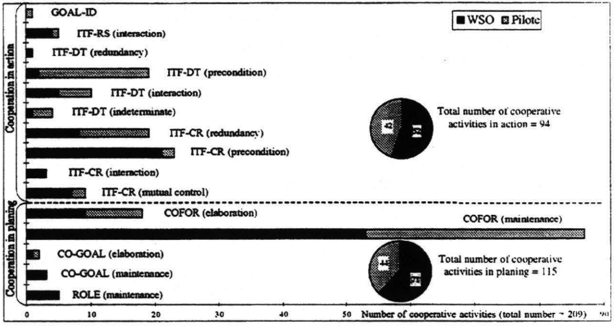


Figure 2: Distribution of the agent in the different cooperative activities

(iii) The coded protocol collected allows us to focus on some sequences of the activity or to analyse its dynamics, which points out some specificity of the cooperative activity:
- The COFOR elaboration episodes mainly originate in the P's individual activity. They are triggered by problems P encounters. Then P commits WSO.
- Such a distribution of cooperative activities is not sufficient, although it is necessary, to reach evaluation and derivation of support design principles. For example, establishing the positive or negative status of interference requires consideration of the dynamics of the activity and evaluation criteria that remain to be defined (a first attempt can be found in Hoc, 1998).

CONCLUSION

This study fulfilled its goal of assessing the relevance of a scheme for coding cooperative activities in the fighter aircraft cockpit and of identifying the nature of these activities. The results can improve the accuracy of the cognitive architecture used to analyse cooperation and thus the structure of the coding scheme. Now we can introduce some important relationships between the two cooperative activity levels: cooperation in action and cooperation in planning. In this way, it seems that representation maintenance activities are at the intersection between cooperation in action and cooperation in planning. We can also abandon too procedural a point of view on cooperation by reintroducing representational structures more explicitly.

In the future, we will devote more attention to the relationships between cooperation and communication. Especially, we will focus on the respective roles of verbal communication within the two cooperative activity levels and the limits of what can be inferred from the sole partner's behaviour observation.

However, this kind of analysis remains labour intensive. We will test some means of reducing the coding scheme to some relevant variables in order to enable us to compare and evaluate cooperation across different conditions with several teams more economically. The main stake is the possible loss of understandability of the protocols and weakening of link to support design principles.

Our results were used to design some assistance for human cooperative activities. The importance of verbal communication leads us to propose dialog windows allowing the exchange of written messages. This assistance will increase the persistence of the information conveyed and avoid the interference of communication on the partner activity. In order to support the distribution of information about the team's activity in the cockpit, several kinds of support mechanisms are proposed: an video picture of the partner and some information about the configuration of the partner work station are displayed on each post, and a chronicle of the major tasks to be done in the cockpit (and their state of evolution). To support the COFOR management, a shared work space (Lemoine-Pacaux & Grislin-Le Strugeon, 1998) allow the operators to put different pieces of information in a common representation.

The continuation of the research program would be dedicated to the study of these assistance and to the validation and development of the present results that, for the moment, depend on a single observation.

ACKNOWLEDGEMENTS

We thank René Amalberti for his scientific advice, the Istres test base staff for its collaboration (especially Isabelle Farsy and Didier Ribot), and Xavier Chalandon and Yves Davaine for reviewing the paper.

REFERENCES

Castelfranchi, C. (1998). Modelling social action for agents. *Artificial Intelligence, 103,* 157-182.

Clark, H.H. (1996). *Using language.* Cambridge, UK: Cambridge University Press.

Hoc, J.M. (1988). *Cognitive psychology of planning.* London: Academic Press.

Hoc, J.M. (1998). How can we evaluate the quality of human-machine cooperation? In F. Darses & P. Zaraté (Eds.), *COOP'98. Third International Conference on the Design of Cooperative systems* (pp. 121-130). Le Chesnay, F: INRIA.

Hoc, J.M., & Amalberti, R. (1995). Diagnosis: some theoretical questions raised by applied research. *Current Psychology of Cognition, 14,* 73-100.

Kræmer, K.L., & Pinsonneault, A. (1990). Technology and groups: assessment of the empirical research. In J. Galegher, R.E. Kraut, & C. Egido (Eds.), *Intellectual teamwork* (pp. 375-405). Hillsdale, NJ: Lawrence Erlbaum Associates

Lemoine-Pacaux, M.P., & Grislin-Le Strugeon, E. (1998). *Multi agent system and human machine cooperation.* Paper presented at the *17th European Annual Conference on Human Decision Making and Manual Control.* Valenciennes, France, Dec.

Newell, A., & Simon, H.A. (1972). *Human problem-solving.* Englewoods-Cliffs, NJ: Prentice-Hall.

Salas, E., Prince, C., Baker, D.P., & Shrestha, L. (1995). Situation awareness in team performance: implications for measurement and training. *Human Factors, 37,* 123-136.

RÉSUMÉ

ÉVALUATION D'UNE MÉTHODE POUR ÉTUDIER LA COOPÉRATION COGNITIVE DANS LE PILOTAGE D'AVION DE COMBAT

Ce texte présente les principaux résultats de l'étude préliminaire d'un projet concernant la coopération entre humains (et son soutien) dans le domaine du pilotage d'avion de combat biplace. A partir des données comportementales et verbales (verbalisations simultanées et en autoconfrontation) obtenues lors de l'observation d'un équipage d'essai sur un simulateur générique et partiel d'avion de combat biplace, une analyse de protocole "individuel" a été réalisée. Cette analyse a permis d'évaluer la pertinence du schème de codage de l'activité coopérative issue du cadre théorique de Hoc (1998), ainsi que de l'ajuster (à la situation cible). Il a permis aussi de mettre en évidence les principales caractéristiques de la coopération dans les cockpits d'avions qui permettront ensuite de proposer une certain nombre d'aides dont l'étude et l'évaluation sera la prochaine étape du projet de recherche.

Mots clés: Coopération cognitive, CSCW, Analyse de Protocoles Individuels, Schème de Codage, Pilotage d'Avion de Combat.

ROLE OF A COMMON FRAME OF REFERENCE IN COGNITIVE COOPERATION
Sharing tasks in Air-Traffic Control

Xavier CARLIER and Jean-Michel HOC
CNRS - UVHC
LAMIH - PERCOTEC
Le Mont Houy - BP 311
F-59304 VALENCIENNES CEDEX
E-mail: Xavier.Carlier@univ-valenciennes.fr
Jean-Michel.Hoc@univ-valenciennes.fr

ABSTRACT

This study deals with cognitive cooperation in the context of the design of a cooperative computer support for aircraft conflict detection and resolution in Air Traffic Control. In order to specify some cooperative capabilities of such a system, we have observed a situation where two radar controllers (RC) had to cooperate for the management of the traffic within a single sector. This paper mainly reports our analysis of the verbal data recorded during a simulated scenario. The results enabled us to describe the elements of a common frame of reference (COFOR), elaborated and updated by the two RCs. They also show the role of this COFOR in the implicit detection and resolution of interference between the RCs' individual activities.

KEY WORDS: Cognitive Cooperation, Protocol Analysis, Air Traffic Control, Common Frame of Reference

INTRODUCTION

This study deals with human-human cooperation in Air-Traffic Control (ATC) as a model to guide the design of human-machine cooperation. After several studies on the same topic within a dynamic task allocation paradigm (Debernard, Vanderhaegen, & Millot, 1992; Hoc & Lemoine, 1998), the current study addresses the question of human-machine negotiation during cooperative problem solving. Cognitive cooperation appeared to be a crucial in these attempts to design artificial agents sharing tasks with the human operators at an acceptable level of performance and reliability (Clark, 1996; Hoc, 1998; Jones & Jasek, 1997; Millot & Lemoine, 1998).

In France, the "en route" air space is divided into sectors controlled by two operators. The Radar Controller (RC) is in radio connection to the aircraft and has a tactical task of guiding the aircraft through the sector. The Planning Controller (PC) is in charge of the coordination between adjacent sectors and of the regulation of the RC's workload. ATC controllers are confronted with a continuous increase in air traffic (about 7% per year). Several solutions to this problem have been implemented or envisioned: Regulation of the clearance for taking off at the European level, Decomposition of sectors into smaller ones, Free flight principles with automatic anti-collision device aboard the aircraft, etc. Various ergonomic approaches are also explored: Representational support (e.g., the ERATO project: Leroux, 1991), Operational support (e.g., SPECTRA using an automatic aircraft conflict detection and resolution device: Lemoine, Debernard, Crevits, & Millot, 1996).

The previous studies have shown that, despite the precautions taken to allocate independent tasks to the human and the artificial agent, it is difficult to manage a dynamic allocation of tasks without creating interference between individual tasks or procedures. For example, if the designer considers that two aircraft conflicts are independent on a technical basis (the two solutions do not create a future conflict between the two deviated aircraft), it can happen that the human's plan is different from the anticipated one and the two solutions (the human's one and the machine's one) can interfere. Thus, negotiations between the two agents at the planning level are necessary to avoid systematic refusals of the machine's solutions by the human. Such refusals result in a stable workload and a failure of the dynamic allocation aim that seeks to reduce workload. The present study was intended to guide the design of a cooperative machine able to perform negotiation with the human operator.

CONTEXT

The study was developed within a multidisciplinary context, integrating supervisory control researchers and cognitive ergonomics psychologists. The former are exploring the possible design of a computer system able to receive a plan concerning one aircraft (e.g, turn to the right), to compute an optimal trajectory (as regards fuel consumption, planned route, possible conflicts with other aircraft, etc.), to return feedback to the human operator in terms of possible problems, and to put back the aircraft on its route as efficiently as possible. Before coming to precise specifications of such a system, psychologists are in charge of anticipating what kinds of benefit can be expected from such an idea and what kinds of problem can be identified.

Among the psychological studies in progress, the one presented here has already produced some interesting results that will be discussed. The purpose of the study was to design an artificial situation where a PC is working with two RCs. The aim was to study human-human negotiation between RCs to produce some recommendations to the supervisory control researchers. With this kind of method, human-human cooperation is taken as a reference model to design human-machine cooperation. Obviously the transfer is not perfect, due to the machine limitations.

THEORETICAL FRAMEWORK

The theoretical framework adopted here is the approach to cognitive cooperation introduced by Hoc (1998). It stresses two minimal conditions for cooperation to develop:

- Each (cooperative) agent strives towards goals and can interfere with others on goals, resources, procedures, etc.
- Each agent tries to detect and manage such interference to make the individual and common activities easier.

The concept of interference is borrowed from studies on planning (Castelfranchi, 1998; Hoc, 1988) and may appear in diverse forms (e.g., precondition relation, when one of the goals can be seen as a subgoal of the other; interaction relation, when the two procedures must be changed; etc.). At first glance, interference can always be considered as negative. In fact, it increases the workload. However, very often interference plays a positive role (e.g, mutual control, when one agent expresses a disagreement on the other's activity and introduces an improvement).

Following this conception, cooperation is not simply considered as a situation (e.g., structural relationships between agents), but also and mainly as a cognitive activity that cannot be seen developed by any agent working alone. Three levels of cooperative activities are considered, in terms of abstraction: cooperation in action, in planning and meta-cooperation.

Cooperation in Action

This class of activities is directly embedded into the action execution level. It integrates interference creation (e.g., mutual control), detection, and resolution (at the action level by local adjustments). It also embodies identification of the other agents' goals, enabling interference management by anticipation. Such an identification can be immediately derived from expertise in the work domain (e.g., if a controller turns an aircraft deviating it from its route, one can infer from domain knowledge that the goal is resolving a conflict).

Cooperation in Planning

These activities develop at the planning level and can improve the performance of the activities belonging to the first class. They contribute to the elaboration and the maintenance of a Common Frame of Reference (COFOR). This complex notion cannot be precisely defined here. It is very close to team situation awareness (Salas, Prince, Baker, & Shrestha, 1995) and to common ground (Clark, 1996). Salas *et al.* (1995) define shared situation awareness as, at least, the shared understanding of a situation among team members at one point in time. It embodies shared representations like contextual ones, agreements on problem or task definition, etc. COFOR is not only used in communication understanding, but also in action coordination. It integrates various elements, such as common goals, common plans, and role allocation between agents.

Meta-cooperation

This level of cooperative activities was not addressed in this study. It can take a very long time to develop, but can considerably improve the cooperative activities at the two previous levels. It integrates the elaboration of compatible representations (the ability to "translate" one's representations into other kinds of representation more compatible with the others' goals), and the elaboration of a model of oneself or of the others. In this study of cooperation between two professional RCs these elements were already available.

METHOD

Subjects

Seven pairs of professional ATC controllers were used as RCs operating together with the same PC.

Experimental Task

A quite full-scale air traffic simulator was used to reproduce a realistic scenario on a sector familiar to the controllers. A high traffic load was defined in order to justify the presence of two RCs (and a possible assistance from a machine). Before the experiment, an aircraft allocation to the controllers was chosen in order to maximise the possible interference between their activities. Namely, each time a new aircraft appeared the controllers were informed of the allocation by a colour. A controller could only give instructions to aircraft under control (allocated to this controller).

Types of Conflict

There were two types of conflict:

- Conflict between several aircraft controlled by the same RC, but the resolution of the conflict may produce another conflict with an aircraft belonging to the other RC.
- Conflict between several aircraft distributed among the two RCs.

In the first case, the resolution is the task of one RC but adjustment of the two RCs' activities is necessary to resolve possible interference. In the second case, the resolution is a common task. RCs have to resolve the conflict and the possible future problems (e.g., if the resolution causes another conflict) together.

Data Collection and Analysis

Three types of data were recorded:

- Actions on the interface (e.g., instructions to aircraft);
- Main events in the traffic (e.g., aircraft entry, strip entry, etc.);
- Spontaneous oral communications between controllers.

The cognitive activities, especially the cooperative ones, were inferred from the data flow and coded using the theoretical framework and the MacSHAPA software (Sanderson, Scott, Jonhson, Mainzer, Watanabe, & James, 1994). The activities are described following a predicate (activity) / argument (specification of the activity) formalism.

RESULTS

First, we will consider the distribution over the two activity levels before considering the distribution over the cooperative activities situated within a given level and illustrating them by some protocol excerpts. These activities were basically inferred from verbal reports.

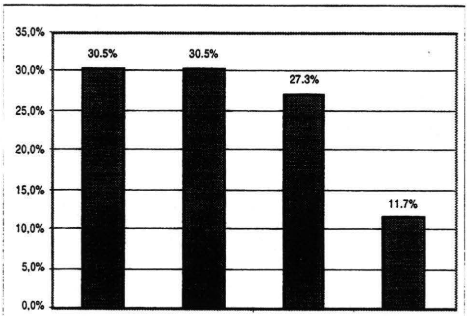


Figure 1: Elementary cooperative activities at the action level

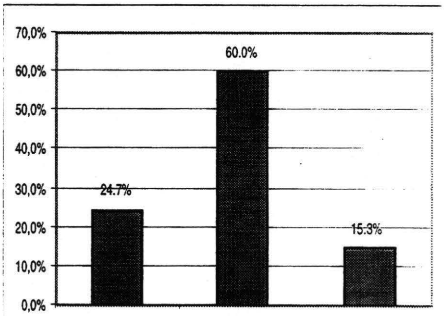


Figure 2: Elementary cooperative activities related to the control activity at the planning level

Levels of Cooperative Activities

On the average, there were 90.1 units per protocol, which overall distribution follows:

- Cooperation in planning: 79.7% ± 10.6%[1]

---

[1] '±' indicates the half width of the confidence interval at the level α=.10.

- Cooperation in action: 20.3%

Cooperation in Action

The distribution of the elementary cooperative activities at the action level is presented in Figure 1:

- Expression of interference by anticipation: 30.5% ± 8.7%
  Ex: " If I turn my AZA to the right, you must turn your AEA to right too."
- interference creation (mutual control): 30.5% ± 11.9%
  Ex: "Are you sure ... You want to turn the AZA behind the MPH ?"
- Interference detection: 27.3 ± 6.8%
  Ex: " Your VIV is behind my DIATC and goes faster."
- Interference resolution : 11.7%
  Ex: " Between my AWD and your HLF, you must turn yours or we will have a problem"

Cooperation in Planning

Two main categories of representation, belonging to COFOR, were reported concerning:

- The process under control (usually considered as shared situation awareness): 36.4% ± 5.0%
- The control activity itself: 63.6%

Common representation of the process under control concerned:

- Group of conflicting aircraft: 61.7% ± 6.9%
  Ex: " AFR1101 and IEA531, they will come across."
  Ex: "The MON is 10 miles behind the MPH"
- Individual aircraft: 38.3%
  Ex: "The BAL is heading 220"

Common representation of the control activity was decomposed into three types (Figure 2):

- Action fulfilment: 24.7% ± 2.6%
  Ex: " I've put my AZA direct to Sauveterre"
- Common goal or plan: 60.0% ± 12.7%
  Ex: "The EIN must go between the IBE and the BAL"
  Ex: "I turn my MON to left and you maintain your HLF"
- Common activity evaluation: 15.3%
  Ex: "The resolution between the BAL and the IBE is almost complete "

DISCUSSION

The allocation of aircraft among the RCs was chosen to maximise possible interference between their activities. The comparison between the two levels of cooperative activities shows the prominence of the planning level. This means that interference between RCs' activities was largely managed at a global coordination level (planning) as opposed to a local interference resolution level.

Not much interference resolution was reported at the action level. This is probably due to the prominence of the planning level in the management of interference and to some automatic activity. Although interference detection is not negligible, interference creation (mutual control) and anticipation, which plays a positive role in cooperation, are well represented at the action level.

The major part of cooperation in planning can be explained by the fact that sharing a large amount of information on the situation (process under control and control activity) can make constraints on RCs' activity more salient. Thus a large proportion of interference detection and resolution is completely implicit (implicitly anticipated). The analysis of cooperation at the planning level reinforces this interpretation:

- At this level we can see that information shared by RCs mainly concerns the control activity (common plan and goal, execution and evaluation of actions) as opposed to the process under control. Thus, interference that could occur at the action level is resolved at this planning level.
- The common representation of the process under control is much more devoted to conflicting aircraft than to individual ones. This means that cooperation in the elaboration of the problem space (representation of the conflicts) plays a major role.

The elaboration of a common goal or plan accounts for a large proportion of the elementary cooperative activities related to the control activity at the planning level. This result may be explained by the fact that many conflicts constituted a common task for the RCs due to the allocation of the conflicting aircraft between the RCs. This type of conflict needs to share a large amount of information concerning plans and goals to be resolved.

The method used to code the activities and the MacSHAPA software did not enable us to directly access the continuous updating of the COFOR during the resolution of a conflict. A more clinical description of verbal protocols is in progress to show the changes in the individual (and collective) problem space resulting from cooperation.

The COFOR has two functions related to the problem space:

1) It permits agents to understand each other's activities.
2) It permits agents to adjust the compatibility of each RCs' individual problem spaces.

Problem space elaboration goes hand in hand with resolution elaboration. For example, a RC may consider the resolution of a conflict between two aircraft and the other RC then identifies that this solution may cause a new conflict with one of the aircraft under her/his control. The first RC must then reconsider the decision, thus it is no longer a two-aircraft conflict but a three-aircraft conflict. With this example we can see the impact of cooperation on the evolution of the individual problem spaces. One RC considers a resolution corresponding to a certain representation of the conflict. This individual communicates a solution regarding this representation. Then, the other RC, who has a different representation of the same problem, can intervene on the first RC's problem representation by introducing a new aircraft that must be considered in the choosing of a new resolution. Thus, the two RCs have now a common representation of the conflict.

CONCLUSION

The results of this study show that the most crucial task in ATC is the elaboration and maintenance of an appropriate problem space. Within this structure the choice of solutions and their interpretations are almost obvious. That is why COFOR elaboration activities are prominent in human-human cooperation when two controllers share the traffic. Certainly, the allocation of schematic solutions and of their implementation to a machine could alleviate RC's workload, especially working memory load since the execution of the operations needed by conflict resolution are distributed over ten or twenty minutes. But the relationships between these solutions and a changing problem space should be carefully considered.

The results presented in this study are useful for the design of the computer system which is being developed through the collaboration between supervisory control researchers and cognitive ergonomics psychologists. The current limitations of machines in terms of COFOR elaboration and maintenance cannot let us hope that the human-human cooperation model suggested by these results could be transferred to human-machine cooperation. However we have suggested that a support for COFOR be designed to serve as a communication medium between RC and the machine. The use of such a medium could be very different for the human and for the machine. RC could use it to spatially define and update the problem space (on an interface similar to radar screen). The machine could use it in a less interpretative way, just displaying information on conflicts unanticipated by RC and gathering information from RC on aircraft allocation and plans.

The COFOR allows the RC to know constraints on the other's activities and then to forecast:

- Interference produced on the other RC's activity;
- Interference resulting from the other RC's activity.

RCs must then integrate interference between individual activities for decision making.

Given the vast amount of implicit information shared in the COFOR, an analysis framework issued from theories of language and especially from pragmatics may be useful for the study of cognitive cooperation.

REFERENCES

Castelfranchi, C. (1998). Modelling social action for agents. *Artificial Intelligence, 103,* 157-182.

Clark, H.H. (1996). *Using language.* Cambridge, MA: Cambridge University Press.

Debernard, S., Vanderhaegen, F., & Millot, P. (1992). An experimental investigation of dynamic allocation of tasks between air traffic controller and A.I. system. Paper presented at the *5th IFAC/IFIP/IFORS/IEA MMS.* Amsterdam.

Hoc, J.M. (1988). *Cognitive psychology of planning.* London: Academic Press.

Hoc, J.M. (1998). How can we evaluate the quality of human-machine cooperation? In F. Darses & P. Zaraté (Eds.), COOP'98. *Third Internationnal Conference on the Design of Cooperatives Systems* (pp. 121-130). Le Chesnay, F: INRIA.

Hoc, J.M., & Lemoine, M.P. (1998). Cognitve evaluation of human-human and human-machine cooperation modes in air traffic control. *International Journal of Aviation Psychologie, 8,* 1-32.

Jones, P.M., & Jasek, C.A. (1997). Intelligent support for activity management (ISAM): an architecture to support distributed supervisory control. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems ans Humans, 27,* 274-288.

Lemoine, M.P., Debernard, S., Crevits, I., & Millot, P. (1996). Cooperation between humans and machines: first results of an experiment with a multi-level cooperative organisation in air traffic control. *Computer Supported Cooperative Work, 5,* 299-321.

Leroux, M. (1991). *ERATO: cognitive engineering applied to air traffic control* (Research Report). Toulouse, F: CENA.

Millot, P., & Lemoine, M.P. (1998). *An attempt for generic concepts toward human-machine cooperation.* Paper presented at *IEEE SMC.* San Diego, CA, Oct.

Salas, E., Prince, C., Baker, P., & Shrestha, L. (1995). Situation awareness in team performance: implications for measurement and training. *Human Factors, 37,* 123-136.

Sanderson, P., Scott, J., Jonhson, T., Mainzer, J., Watanabe, L., & James, J. (1994). MacSHAPA and the enterprise of exploratory sequential data analysis (ESDA). *International Journal of Human-Computer Studies, 41,* 633-681.

RÉSUMÉ

RÔLE D'UN RÉFÉRENTIEL COMMUN DANS LA COOPÉRATION COGNITIVE: LE PARTAGE DES TÂCHES DANS LE CONTRÔLE DU TRAFIC AÉRIEN

Cette étude traite de la coopération homme-homme comme un modèle pour la conception d'une situation de coopération homme-machine dans le contrôle de trafic aérien. Il s'agit de développer un système d'aide à la résolution de conflit. Ce système devra être capable de recevoir un plan concernant un avion (par exemple: tourner à droite), et de gérer la trajectoire de cet avion à l'intérieur du secteur le plus efficacement possible, tout en informant en retour l'opérateur en ce qui concerne les éventuels problèmes rencontrés. Afin de retirer un certain nombre d'informations concernant les éventuels bénéfices et problèmes pouvant découler d'une telle idée, nous avons procédé à une étude de la coopération entre deux contrôleurs. Pour cela il nous a fallu créer une situation artificielle sur simulateur où, pour un même secteur, deux contrôleurs-radar (CR) sont chargés de guider les avions traversant le secteur sous contrôle. Chaque CR ne peut intervenir que sur une partie des avions présents dans le secteur. Les verbalisations recueillies ont alors fait l'objet d'une analyse portant plus spécifiquement sur les activités de coopération. Sur un plan théorique, nous avons considéré la coopération cognitive comme intervenant lorsque chaque agent de la coopération poursuit un but et que les activitiés mises en œuvre pour atteindre le but de chacun des agents nécessitent des ajustements lorsque ces activités interfèrent entre-elles. Les activités coopératives qui découlent d'une telle situation peuvent alors être réparties et analysées en deux niveaux. Premièrement, le niveau de l'action où figure la détection, la résolution (par ajustement local), l'anticipation et la création d'interférences (contrôle mutuel). Deuxièmement, on a considéré le niveau de la planification où figure l'élaboration d'un référentiel commun qui consiste en un partage d'informations sur, d'une part, le déroulement du processus contrôlé et, d'autre part, sur l'activité de contrôle elle-même. Une analyse des différentes informations échangées à ces différents niveaux nous indique qu'un grand nombre des interférences entre activités sont gérées non pas localement dans l'action mais à un niveau plus global grâce à l'entretien d'un référentiel commun qui rend plus saillants les interférences entre activités ce qui rend leur gestion pour une bonne part implicite.

MOTS CLÉS : Coopération cognitive, Analyse de protocole, Contrôle de trafic aérien, Référentiel commun

Specifying artificial cooperative agents through a synthesis of several models of cooperation

E. GRISLIN-LE STRUGEON, P. MILLOT
LAMIH CNRS 8530
Université de Valenciennes et du Hainaut-Cambrésis
BP 311 Le Mont-Houy
59304 VALENCIENNES-Cédex, FRANCE
e-mail :{millot, strugeon}@univ-valenciennes.fr

ABSTRACT

Our study aims at designing Decision Support Systems for supervision of complex dynamic processes. This paper proposes primitives, which must be added to artificial agents in order to make these agents cooperative. Our method starts with a definition of cooperation from Cognitive Psychology and derives it by simulating the agents' activities when cooperating. These primitives are grouped in two main classes: Managing Interferences between the agents' goals (MI), and Facilitating the agents' Goals (FG). In order to be exhaustive, we assume that the 3 generic cooperative forms by Schmidt (1991) can be combined to describe any cooperative situation. We then discuss the technical feasibility of these primitives in the light of present issues in the field of Artificial Intelligence.

Keywords Human-Machine cooperation, Know-How, Know-How-To-Cooperate, artificial agents

INTRODUCTION

The applicative field of our study is the supervision and control of large, complex dynamic systems (eg : Air Traffic, Power plants ...) in which the nature of the interactions between humans and Decision Support Systems (DSS) is directed toward Human-Machine cooperation. Previous papers have distinguished three main concepts 1) the structure of the organisation in which DSS and humans interact, 2) the know-how (KH) of the DSS which includes its problem solving capabilities (knowledge and processing capabilities) and communication means (with the environment and the other agents), and 3) specific capabilities needed to cooperate with other agents, called Know-How-to-Cooperate (KHC) (Millot & Hoc, 1997).

The complementary interests of structural (hierarchical versus heterarchical organisation) and functional models of cooperative systems have been debated (Millot & Lemoine, 1998, Millot, 1998). Using a multidisciplinary approach which regroups mainly "human-engineering models" and cognitive psychology approaches, we have proposed rough specifications of primitives for designing the KHC of an artificial agent (Millot, 1998). The present paper aims to analyse these primitives and their technical feasibility for an artificial cooperative agent, in the light of Distributed Artificial Intelligence approaches.

Our study does not pretend to cover all kinds of cooperation especially between humans or even between humans via machines (CSCW). Our contribution is in the human engineering field, and aims to provide guide-lines for an engineer who designs concrete and usable cooperative decision support systems.

PRELIMINARY SPECIFICATIONS FOR COOPERATIVE AGENTS

An Attempt to Define the Know-How-To-Cooperate capabilities

In the field of cognitive psychology, Hoc (1996) and Millot & Hoc (1997) propose the following definition : "two agents are cooperating if 1) each one strives towards goals and can interfere with the other, and 2) each agent tries to detect and process such interference to make the other's activities easier". Two classes of activities dedicated to cooperation and constituting the know-how-to-cooperate (KHC) capability (Millot & Lemoine, 1998) can thus be derived :

- the first class requires capabilities for detecting and managing interferences between goals (coted MI : Managing Interferences) : interferences can be seen as positive ones (common goal, subgoal ...) or negative (conflicts between goals, subgoals ... or common resources to be shared).
- the second class requires capabilities for making other agents' goals easier (coted FG : Facilitating Goals).

According to these classifications MI includes more management capabilities, while FG concerns a more benevolent behaviour by the agents. Given an agent AGx that includes a know-how KHx, and a know-how-to-cooperate KHCx, and which is within a structure, our purpose is to try to specify KHCx through MIx and FGx. With this in mind, we simulate (predict) the activities of AGx when cooperating with AGy in the different cooperative situations which can be encountered (or built). These different situations result from the generic typology of cooperative forms proposed by Schmidt (1991): augmentative, debative, integrative (Millot, 1998). Our hypothesis is that any cooperation situation may be described as a combination of the three forms mentioned above[1]. We will then define the agent's KHC for each cooperative form. Designing the KHC capabilities of an agent for dealing with any cooperative situation will then result in combining the KHC capabilities related to each form. This approach appears more deductive (though not contradictory) than Castelfranchi's one (1998) which starts from a typology of the interferences and proposes cooperative activities for dealing with.

Cooperative forms and Know-How-to-Cooperate capabilities

Augmentative Cooperation

Cooperation is **augmentative** when agents have a **similar know-how** but the agents must be multiplied to perform a task too demanding for only one agent ; the task T is then shared into **similar subtasks STi**. Interferences between the agents' activities can result from sharing common resources, but also from conflicts (concerning goals or subgoals) between their own STi. The KHC activities are: 1) decompose the task T, before it is performed, into **independant** STi in order to prevent these conflicts, 2) manage the residual conflicts (for instance concerning common resources) during the STi execution, and 3) recompose the results afterwards if the results of the STi were not actions, or else rebuild the task context. These activities concern especially MI capabilities and we assimilate them with **coordination activities**. They can be performed by a third agent called a **coordinator**, or by one of either AGx or AGy, which then plays a double role of coordinator and actor.
The coordinator's KH regroups capabilities for acquiring the task context, for building a global plan for the task and for decomposing it into STi (sub-plans).The coordinator's KHC includes capabilities for acquiring other agents' KH (or infering it through a model of them), acquiring other agents' constraints (eg : workload WL) and/or resources for the sub-task allocation, controlling and recomposing the results or the contexts after each STi has been performed, and managing conflicts on common resources. All these KHC capabilities concern MI. The other agents' KHCs are FG and consist of answering the coordinator requests.
Indeed, introducing a coordinator is a choice which strongly links the functional model with the structural aspect of the multi-agent organisation, especially the hierarchical structure. The example given in Table 1 follows this choice, in which MI capabilities of AGx allow it to be a coordinator. Furthermore FGx capabilities allow it to be controlled by another coordinator. However, we can imagine a more general organisation in which the coordination could be shared between both agents through a higher level cooperation. But as we will see in the next section, several feasibility constraints particularly related to the limited negociation capabilities of the artificial agents make this option difficult to implement.

Table 1 : Summary of KHC activities of the agent AGx in an augmentative cooperation with AGy

		MIx	FGx
Before	Decomposing T into STi	acquiring KHy	transmitting its KHx to AGy
	Allocating STi	- acquiring WLy - allocating STi to AGy and to itself	- transmitting its WLx to AGy - acknowledgement of its intention to perform STi
Along	Managing conflicts between AGx and AGy on common resources	- detecting conflicts with AGy - acquiring conflict signals from AGy - affecting priorities to AGy and to itself - affecting resources to the other afterwards	- informing AGy of the occurrence of a conflict - informing AGy that the resource is available
After	Recomposing / controlling the results	- acquiring results (context) from AGy - controlling results (context) from AGy	- transmitting results (context) to AGy - transmitting sub-plans (sub-goals) to AGy

---

[1] This hypothesis is not yet formally proved, but several examples make it coherent, and a future paper will study it in further details.

Debative Cooperation

Cooperation is **debative** when agents have **a similar know-how** and are faced with **a unique task** T (non-shareable into STi). Each agent solves the task and then debates on the results with the other agent. Conflicts can appear and the KHC must allow these conflicts to be solved by the means of explanations based, for instance, on previous partial results in the problem solving pathway and on a common frame of reference (Millot, 1998). Before task execution, each agent KH consists of the capabilities to acquire the task context, and to build a plan (goal, sub-goal ... means), (see Table 2). Each agent KHC consists of acquiring the other agent KH either by infering a model of the other agent, or by asking the other agent for information (MI capabilities). The answer to this request by the other agent is FG.

After the task execution (or partial execution), each agent transmits its own results to the other, receives the result from other and compares these results with its own. In addition to MI capabilities (asking for results from the other) and FG capabilities (transmitting its own results), this activity requires specific skills in order to understand the other's results, to compare with its own and to decide whether it agrees with them or not. All these skills are MI.

In case of conflict, each agent must be able to ask for explanations (the other agent's view on the task context, its partial results, its goal, sub-goals ....) to compare these explanations with its own view-point and to make a decision as to whether the conflict should persist or not. In addition, each agent must be able to accept its errors and to "learn from it" for instance by modifying its own plan or in requiring the other agent to do so. As in the augmentative form, MI capabilities can be assimilated with **coordination** activities. This seems realistic, as the unique task can be controlled by one agent which then becomes the "natural" coordinator of the debate.

Table 2 : Summary of the KHC activities of the agent AGx in a debative cooperation with AGy

	MIx	FGx
Preparing cooperation	acquiring (infering) KHy	transmitting its KH to AGy
Comparing results (even partial)	- acquiring results from AGy - understanding results from AGy - comparing with its own results, - deciding : agree, disagree	transmitting results to AGy
Conflict solving	- acquiring AGy's plan and view- point on task context - comparing respective contexts and plans - deciding on the conflict persistence and asking AGy (or itself) to modify its plan	- transmitting its own plan and view-point on the context to the other - accepting its own error and "learning from it" (modifying its own plan)

Integrative Cooperation

Cooperation is **integrative** when agents have **different and complementary know-how**, and the task T can be shared into **complementary sub-tasks STi** related to each KH. As in the previous forms, let us consider the role of **coordinator** being performed by one of the agents (AGx or AGy) or by a third agent.

The coordinator's KHC must 1) elaborate a common plan (goal, means) and decompose it into complementary sub-plans (STi, sub-goals) each related to the respective agents' KH, 2) control the partial results or the context evolution during the execution of the STi by the agents, and 3) recompose the results afterwards if the results of the STi were not actions, or else rebuild the task context. There is a similarity with the augmentative form, except that during the STi executions, due to the possible interactions between agents, for which capabilities for checking each partial result and for ordering corrections if needed are required.

Table 3 summarises MI activities which are more related to coordination capabilities, and FG activities which are more concerned with a benevolent behavior to satisfy coordinator requests. This example considers that the shared plan has already been elaborated by the coordinator. Building the shared plan is another cooperative activity which can be related to a debative form (see Millot, 1998).

These cooperation mechanisms roughly described above can be derived into primitives for building the algorithms which must be added to the artificial agent or the relevant interfaces for the human agents. The main idea is to provide the artificial agents with a relevant combination of the primitives related to each cooperative form so that it be able to cooperate in any situation.

Table 3 : Summary of the KHC activities of the agent AGx in an integrative cooperation with Agy: case of a shared plan elaborated by a coordinator

		MIx	FGx
Before	Decomposing T into STi	acquiring (inferring) KHx	transmitting its KHx to AGy
	Allocating STi	- acquiring WLy - allocating STi to AGy and to itself	- transmitting its WLx to AGy - acknowledgement of its intention to perform STi
Along	Controlling partial results	- acquiring partial results (context) - comparing results to plan and to sub-plans - ordering corrections to AGy or to itself if needed	- transmitting (partial/results context) to AGy - correcting the STi if asked by AGy
After	Recomposing / Controlling the results	- acquiring results (context) - controlling results (context) - ordering corrections to AGy or to itself if needed	- transmitting results (context) to AGy - correcting the STi if asked by AGy

COOPERATION CONDITIONS AND MECHANISMS

Our aim is the design of cooperative artificial agents with a view to implement human-machine and machine-machine cooperative organisations for the supervision of dynamic systems. Let us examine now the technical feasibility of the principles proposed above through methods presently well established in Artificial Intelligence, specifically in the field of Multi-Agent Systems (MAS).

The Know-How-to-Cooperate of the Artificial Agents

In MAS, an artificial agent can have three functional levels, at the lowest level, representation is bound to perception-action rules ; at the medium level, KH is needed for solving problems in an individual way ; at the highest level KH and KHC are dedicated to the cooperation with other agents and can be seen as the background for social behavior in the organisation. In the so called "reactive agents" (like the ant nest model), the medium level has quite poor cognitive abilities, while this level is more achieved for "cognitive agents". A classic DSS can then become a cognitive agent if its third level KHC is completed. For that purpose the main functions of the two upper levels must include a model of the agent environment, of the other agents (including their goals and KH) and of itself (its skills and abilities). All of its actions and decisions are made according to this knowledge about itself and the others. Moreover, a representation of the group or of the organisation in which it participates directs the actions of the agent: its behaviour is controlled by a set of rules, as in any society. These upper levels allow the cognitive agent to play a role in decisional autonomy.

These individual and social representations constitute the "shared system of reference"of the agents (Lemoine-Pacaux & Grislin-Le Strugeon, 1998). It establishes the bases on which the agents define several procedures in order to cooperate. With this knowledge, they develop KHC abilities. The following paragraphs describe in which way the KHC primitives are (or could be) implemented and the difficulties that are encountered.

Before task execution

- *to acquire KHy, WLy, goals and plans of AGy, AGy's "profile" (skills, abilities, knowledge, know-how)* : The most elaborate technique consists of modelling AGy by inference, from the knowledge AGx has on Agy and from the observation of AGy's actions. However, this method is complex, slow and prone to problems due to possible inconsistencies between the model and observed reality. Thus a currently used technique consists of building the agent model by asking to it the needed information (information exchanges) as suggested in the tables above. Such a model interprets the intentions of the other agents as presented in Geddes (1994).
- *to elaborate a global plan, to divide it into sub-plans and to allocate STi to the agents.* The joint plan must allow allocating the different tasks to the agents. Elaborating a plan presupposes decomposing a task T into STi in order to reveal the possible interferences between STi (compatibility), material resources and "agents" resources in terms of skills, abilities, and availability. The distribution is determined from the knowledge they have about each other.

  Distributed planning methods exist ((Mandiau & Piechowiak, 1998), (Mouaddib, 1997)), involving all the agents in the elaboration of the joint plan. These methods are based either on partial plan exchanges (parts of plans arranged to be integrated in a joint structure), or on the iterative construction of the global plan, or on the choice of one global plan from those proposed by the agents.

However, the most efficient method consists of entrusting planning to a unique **coordinator** agent. In this case, the planning agent must possess the most complete representation possible of the existing resources (material and agents). This completeness being rarely achieved, adjustments are necessary during the realisation. Accepting the joint plan can imply either the withdrawal of individual goals or their integration into the joint plan. To achieving that, agent behaviour can be controlled by the following principles :
- the individual actions maximise the results (or the efficiency) of everyone. For instance, if AGx believes that achieving AGy's goal is a way to achieve its own, its goal will be that AGy achieves its goal.
- the agents are mutually committed on a joint plan or a social exchange. For instance AGx does something for AGy so that AGy will do something for AGx (see, for example, Sichman, 1995). Moreover, the acceptance of task responsibility can be materialised, in some systems (Sen & Durfee, 1994), by a "commitment" to achieve its realisation.

- *to ask for information to verify or complete knowledge.* In anticipation, this activity aims either to reinforce the "shared system of reference", or to improve the knowledge about the other members of the group (goals, plans, knowledge, abilities, know-how) which is necessary for cooperation.
- *to transmit information (its KH, its WL, its intention to perform STi, the availability of a resource).* The most cooperative behaviour consists of taking the initiative to give a piece of information considered to be useful to other agents. This presupposes that the agent is able to infer the other agents' needs, from the knowledge about them that it has at its disposal. Such a behaviour comes from a corresponding goal or a rule such as: "if AGx believes that AGy needs to know the fact F, and AGx knows F, then AGx informs AGy about F." More simply, the agents can broadcast their skills, abilities and goals inside the group in order to allow all group members to build their mutual representations.

During the task execution

- *to refine the knowledge of the other agents,* based on the observation of the other agents actions and on the exchanged information. In real time, it happens that agents ask for information when facts are different from what was expected. Some systems use the notion of "trust", equivalent to a level of belief in the information provided by others (trust granted or not to a given agent according to its supposed competence and knowledge), to solve possible inconsistency problems in the knowledge bases (Huhns & Bridgelan, 1991).
- *to transmit results, to point out the occurrence of a conflict.* This is realised at the individual level in performing the allocated tasks, including possible coordination actions (transmitting results, waiting for information...). Even if the complete knowledge of the joint plan by everyone is rarely achieved, partial knowledge allows to manage local interferences.
- *to acquire information (conflict signal, results)* and *to reason (comparing plans, results)* in order to adapt behaviour to the situation: *affecting priorities to resources, ordering corrections.* For example, integrating the others' actions in order to benefit from the modifications made on the environment.
- *to correct the STi if asked by another agent...* even if it means suspending the present activity. The difficulty lies in the management of priorities between individual and social goals. A social goal is relevant to the interest of the group and the individual goal only contributes to one individual's interest. But a social goal can usually be divided up into sub-goals which are dispatched among the agents. To which extend are these sub-goals integrated as individual goals (loosing consequently their social goal status)? And how can the priority between two social sub-goals be managed ? This priority management is rather ill-defined in MAS. A current approach consists in evaluating the goal priority according to the respective hierarchical positions of the concerned agent: in other words, the need expressed by the hierarchical superior prevails over the individual goal. In a more complex organisation, a possibility is to negotiate in order to persuade the other agent to withdraw its individual goal to the benefit of another one. A rather well-advanced protocol described in Chang & Woo (1992), is based on a set of speech act verbs classified and constrained by temporal relations. For example, "DISSENT" must appear after "MAKE CLAIM", which itself is followed by an alternative "AGREE" or "DISSENT".

After the task execution

- *to transmit information: (its view on the context, its results), to acquire and control results* in order to compare with what was expected, *to order corrections if needed, to correct the STi if asked by another agent.* (for these 4 primitives see above).
- *to accept its own error and "learn from it".* Knowledge management problems appear here again, as when getting information from other agents: some methods exist but they are difficult to implement effectively and are thus rarely chosen when programming the final system.

CONCLUSION

As shown in this paper the technical feasibility of the primitives which constitute the artificial agents' KHC is far from being achieved. Particularly, major functions, such as elaborating a shared plan, pointing out the occurrence of a conflict between the agents' goals and the others' goals, managing the priorities between the respective social and individual agents' goals remain difficult to implement. Because, the present technical solutions to these points are not fully developped we are forced to fall back on the structural organisation for instance using a coordinator at the higher level of a hierarchy. More generally this shows that a single functional approach for designing a cooperative organisation involving artificial agents is not sufficient. Furthermore, when dealing with human-machine cooperation, complementary human constraints and capabilities must be taken into account. Therefore, this study must be seen as only one technical step for designing human-machine cooperative systems.

REFERENCES

Castelfranchi, C. (1998). Modeling social actions for AI agents. *Artificial Intelligence, 103*, 157-182.

Chang, M.K., & Woo, C.C. (1992). SANP: a Communication Level Protocol for Negotiations. In E. Werner & Y. Demazeau (Eds.), *Decentralized A.I. 3* (pp. 31-54). North-Holland.

Geddes, N. (1994). *A model for intent interpretation for multiple agents with conflicts*. Paper presented at *the IEEE International conference on Systems, Man and Cybernetics*, San Antonio, Texas, Oct.

Hoc, J.M. (1996) *Supervision et contrôle de processus, La cognition en situation dynamique*. Presses Universitaires de Grenoble.

Huhns, M.N., & Bridgelan, D.M. (1991). Multiagent Truth Maintenance. *IEEE Transactions on Systems, Man and Cybernetics, 21, 6*, 1437-1445.

Lemoine-Pacaux, M.P., & Grislin-Le Strugeon, E. (1998). *Multi-agents system and human-machine co-operation*. Paper presented in the *17th European Annual Conference on Human Decision Making and Manual Control*, Valenciennes, France, December.

Mandiau, R., & Piechowiak, S. (1998). *Conflict Solving in Multi-agent Distributed Planning*. Paper presented at the *IEEE SMC*, La Jolla, CA, USA, October.

Millot, P., & Hoc, J.M. (1997). *Human-Machine Cooperation : Metaphor or possible reality?* Paper presented at the *European Conference on Cognitive Sciences ECCS'97*, Manchester UK, April.

Millot, P. , & Lemoine, M.P.(1998). *An Attempt for generic concepts toward Human-Machine Cooperation*. Paper presented at *IEEE SMC'98*, San Diego, USA, October.

Millot, P. (1998). *Concepts and limits for Human-Machine Cooperation*. Paper presented at *IEEE SMC CESA'98 Conference*, Hammamet, Tunisia, April.

Mouaddib, A. (1997). Progressive negotiation for time-constrained autonomous agents. In the proceedings of *the 1st International Conference on Autonomous Agents (Agents'97)* (pp. 8-15). ACM Press.

Schmidt, K. (1991). Cooperative Work : a conceptual framework. In J. Rasmussen, B. Brehmer, & J. Leplat (Eds.), *Distributed decision-making : cognitive models for cooperative work.* (pp. 75-110). Chichester, UK: John Willey and Sons.

Sen, S., & Durfee, E. (1994). The role of commitment in cooperative negotiation. *International Journal on Cooperative Information Systems, 3 (1)*, 67-81.

Sichman, S. (1995). *Du raisonnement social chez les agents.* Doctoral dissertation, Institut National Polytechnique de Grenoble, France.

RESUME

SPECIFICATION D'AGENTS ARTIFICIELS COOPERATIFS A PARTIR D'UNE SYNTHESE DE PLUSIEURS MODELES DE COOPERATION

Notre étude vise la conception d'outils d'aide à la décision pour la supervision de procédés dynamiques complexes. Ce papier propose des primitives dont les agents artificiels doivent être dotés pour être coopératifs. La méthode part d'une définition de la coopération issue de la Psychologie Cognitive et la décline en simulant les activités de coopération des agents pour définir ces primitives informatiques. Ces simulations sont menées dans chacune des 3 formes coopératives de Schmidt(1991) selon l'hypothèse que les combinaisons permettent de décrire toutes situations de coopération. Nous analysons ensuite les faisabilités techniques de ces primitives au regard des avancées actuelles en Intelligence Artificielle.

Session 4: Human-human cooperation
Coopération homme-homme

Analysing Work Activity for the Assessment of Procedure Use in Simulated Control Rooms

The case of emergency operation of nuclear power plants

Geneviève Filippi[*],
Geneviève Saliou,

EDF, Research and Development Division
Safety and Reliability Branch
1 avenue du Général de Gaulle,
92141 Clamart
France

genevieve.filippi@edf.fr
genevieve.saliou@edf.fr

Abstract : The use of simulators is essential to study human factors in emergency conditions of nuclear power plants. In order to ensure efficient coupling between technical arrangements—particularly procedures—and nuclear power plant operating crews in emergency conditions, EDF has set up an 'observatory': a monitoring system for the collection of simulator data. This paper focuses on a data-collection methodology called ANACONDAS. It will highlight the interest of qualitative analysis to explain the use of procedures in emergency operation on simulators.

CONTEXT OF SIMULATOR USE BY EDF: A NUCLEAR POWER PLANT EMERGENCY-OPERATION OBSERVATORY

In order to ensure efficient interlinking between technical arrangements—including procedures—and nuclear power plant control crews in disturbed and accident conditions, EDF has set up a permanent system for collection of simulator data. This so-called 'observatory' of accident operation combines extensive quantitative analyses based on the data gathered by instructors during training sessions, and intensive qualitative data based on tests especially devoted to assessing operators cognitive processes.

This observatory has effects on enhancement of procedures and training, and assessment of choices relating to work organisation. The results also help provide data for updating PHRA.

In this paper we will focus on ANACONDAS, a data-collection methodology for the intensive-analysis part of the observatory, and we will illustrate the value of developing qualitative analysis of emergency operation on simulators for the understanding of procedure use.

---

[*] Corresponding author: EDF, Direction Etudes et Recherches /Etudes de Sûreté et de Fiabilité, 1 avenue du Général de Gaulle, 92141 Clamart, France

Simulators are often considered to be the equivalent of laboratory situations, in that simplification of reality makes it possible to control variables and to reproduce precise experimental conditions for comparison and interpretation of the results by statistical means. This application of simulators is useful when one wants to compare the performance of different crews or to identify and quantify errors.

New topics relating to accident operating and procedure use are emerging; they concern, for example, the evolution of operators knowledge and skills, the role of procedures as a support for reasoning, the modalities for co-operation within the crew and the role of procedures as a support for co-operation, etc. To address these topics, it is necessary to study simulator activity in depth so as to understand how operators accord meaning to a situation and construct their point of view during the recovery from incidents. Questions should also be asked about the collective dimension of the work: how do the technical arrangements in the control room—including procedures—help or sometimes hinder collaboration?

Bearing this in mind, it is possible to obtain revealing results by considering simulator tests not as experimental protocols but as specific working situations in which certain characteristics must be taken into account as such (especially interactions between observers and the operating crew).

However, giving answers to the above questions presupposes that one or several "theoretical object" have been clearly defined. In other words, human activity in nuclear control room settings is too complex to be directly appreciated, it is therefore inevitable to determine the aspects of activity that we choose to study, e.g. (1) the dynamic nature of reasoning linked to the whole of human activity, (2) the importance of the context in which actions take place, or (3) the collective dimension of activity.

For analysing work activity in the simulated control room, we take inspiration from Theureau & Pinsky 's (1990, 1992) conceptualisation of work analysis based on the “course of action”, i.e. the part of activity that is meaningful for the agent. With this approach it is possible to report on individual activities and the interactions that an agent has with the others.

To approach the collective aspects of work in the control room, it is fruitful to combine this kind of approach with other theoretical and methodological frameworks, for example those proposed by :

- recent development of ethnomethodology (Suchman, 1991, Goodwin, 1991; Heath, 1993) and "situated cognition' which stress the joint construction of interaction by partners and the process of mutual adjustment of action,
- cognitive anthropology (Hutchins, 1996) which states that the pertinent level of analysis is not the individual activity but he collective itself within whose framework, cognition is socially distributed between agents and technical artefacts.

ANACONDAS : DATA COLLECTION AND ANALYSIS METHODOLOGY

Data Collection

A full-scale simulator replicates an actual control room; the physical simulation of phenomena is as realistic as possible. A crew consisting of a reactor operator, a turbine operator, a supervisor, and an operations manager takes part in the tests. In certain tests, auxiliary operators also help simulate the interactions between the control room and the field. Accident scenarios are defined beforehand and are enacted several times with different crews. These scenarios often comprise accumulated failures of varying degrees of complexity.

To better situate the context of the methodology, it is necessary to make a distinction between two sorts of activity description:

A. one — **intrinsic description** — represents an internal view of cognition from the point of view of the agent: it addresses cognition, i.e. reasoning, actions, and communications “here and now”.

B. the other — **extrinsic description** — allows for characterisation of activity on the basis of its external factors, from the point of view of an observer, e.g. the different aspects of the situation such as the state of the process simulated and its evolution during the test, the performance of the crews (determined by execution or non-execution of certain important actions, and the time taken to execute them), the level of agent training, their culture, etc.

This distinction between intrinsic and extrinsic description is fundamental in our approach: we postulate that cognitive activity can be analysed through an intrinsic description in the light of extrinsic descriptions.

A. The *intrinsic description of activity* is based on the combination of data from observation of activity during the test and from induced verbalisation after the test. The methodology aims at restricting analysts' interpretations as much as possible so as to take account of the phenomena inherent to the activity.

- data on behaviour during the test is gathered by observers (who take notes) and by video and audio recordings (two cameras) of actions and communications. Recordings are continuous in order to take account of the dynamics of activity.

- induced verbalisation data is gathered in self-confrontation interviews: the analyst presents the video tape of the test to the operator so that the operator can explain his actions and communications and give his interpretation of events. The purpose of these verbalisations is to allow the agent to reconstitute his course of action in context.

Transcription of this observation data and of the verbalisations provides a script that records the temporal organisation of actions. It is the basis for analysis of the data.

B. Data for *extrinsic description of activity* is quite varied. It essentially concerns:

- .information gathered during debriefing meetings with the crews who participated in the tests for better characterising their training and their experience in the job,
- the judgement of instructors and technical experts on how the scenarios were handled—in terms of technical performance and crew behaviour—based on their observation of the tests,
- all other sources of data, particularly those concerning written evidence such as the computer log of the test, documents filled out by operators during the test, etc.

Data Analysis

With a view to carrying out more thorough studies of emergency-operation phenomena, we intend to work towards modelling the individual and collective activity of the operators in the simulated control room. This implies analysing operator activity as a whole in order to both describe the dynamic organisation of actions and to explain this organisation by reconstituting the underlying reasoning.

This sort of analysis differs from error-based behavioural analyses of operators. Like many authors (Winnograd & Florès, 1986, Bannon, 1991, Carrol, 1997), we consider that the agent's point of view is preferable when it comes to interpreting his activity.

Analysis of emergency operation involves firstly taking account of how each agent construes the global organization of his activity, and secondly identifying regularities in several agents' activity , in other words, modeling the organisation of this work activity. Let us consider these two steps of data analysis.

The construction of reduced accounts. First of all, this means dividing the course of the activity into units that are meaningful for the agent, and naming them by answering the question : " from the agents point of view, what is this about?" These meaningful units reflect the fact that each action is not isolated but included in sets of temporally organised actions that are coherent for a given agent. Each script divided into meaningful units in this way produces a reduced account and provides a particular description of the scenario observed.

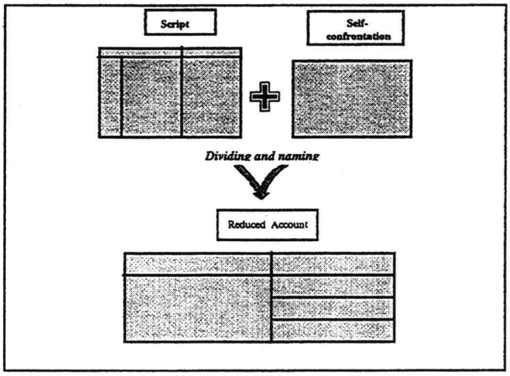


Figure 1: The construction of the Reduced Account

A model of the global organisation of the activity corresponds to generalisation based on comparison of the reduced accounts. This process of comparison and generalisation of reduced accounts reveals archetypal sequences reflecting regularities amongst agents and amongst crews in operation activity on simulators.

It is the identification of these archetypal sequences that makes it possible to compare the agents logic with the logic underlying the design of the technical set-up. It thus becomes possible to examine procedure-design principles with respect to the logic of reasoning and to detect any discrepancies.

WORK ACTIVITY LOGIC AND PROCEDURE-DESIGN LOGIC

In order to show the interest of such an approach, let us consider the topic of procedure use in emergency operation on simulators. In incident or emergency situation, the operators of French nuclear power plants use operation documents presented in the form of logic-diagrams. According to the answers given to different tests relative to the state of the process variables, the operators are supposed to follow step by step specific paths in the procedure where they are asked to carry out actions to bring the nuclear power unit towards off-position enabling normal condition recovery.

The tasks predefined by procedure designers do not claim to represent the steps of human reasoning. On the contrary, analysis of the managing of a simulated nuclear power unit in emergency condition highlights the fact that agents do not merely apply procedures, but above all, they solve problems showing their active engagement in the situation.

The systematic analysis of activity on simulators enables to emphasise the part played by initiative required for the procedures to be applied efficiently in the particular context of each test. It has been noted, thus, that less than half the significant units of a test (the significant units are not proportional in time) correspond to the strict application of the procedure. The other segments of the activity correspond to a permanent task of overall understanding of the situation in order to identify changes and monitor the progression of the test, ensure that the actions requested by the procedures match with the state of the process, interpret the procedures themselves, etc..

Figure 2 represents the reduced account of a short fragment of the transcription of the Supervisor's activity during a test. It represents different levels of activities' structure in which the participants are engaged. The finest level is that of the elementary significant units which records the organisation of actions and communications as they emerge step by step. The second level is that of the sequences in which groups of significant units are embedded. The last level is that of the macro-sequences expressing actor's engagement at a higher level. Each level of this structure can be read autonomously and renders a more or less abstract account of the script .

In this extract, the first sequence is relative to the monitoring of level II operating instruction. The Supervisor is aware that the situation requires the application of the level III instruction, because one of the steam generator (SG) is radioactive. However, he can not take this level III instruction until the turbine operator has finished isolating the SG and tick off the state of this SG in a specific box of a crew reference document. At the same time, the Supervisor has noticed that, shortly before and thanks to its procedure that the AFS is no longer functioning; he must at the same time solve this specific problem (macro-sequence "handling the AFS loss") and integrate it in the rest of the operation (sequence "ensuring the request for cooling is appropriate with the AFS loss").

The Supervisor's course of action		
Elementary significant units (from the point of view of the Sup, what is happening here?)	**Sequences**	**Macro sequences**
1. Asks the Reactor Operator if he has requested a stabilization of the primary coolant temperature →no, it is cooling at 56°C/h	Applying a level II operating instruction while waiting to be directed towards a level III instruction	**Applying an instruction not adapted to the situation while waiting for the adequate ticking off of the steam generators**
2. Is surprised that there has been a cooling requested when there is no auxiliary feedwater supply	Ensuring that the request for cooling is appropriate with the auxiliary feedwater supply loss	**Handling the auxiliary feedwater supply loss**
3. Applies level II operating instruction sequence 4b [monitoring of Steam Generator]	Applying level II instruction while waiting to be directed towards a level III instruction	**Applying an instruction not adapted to the situation while waiting for the adequate ticking off of the steam generators**
4. Reactor Operator asks him if he too is starting again sequence 4 → no, because he has a problem on the auxiliary feedwater supply		
5. Applies level II p.4b instruction [monitoring of SG] → nothing in this module orients him towards level III operating instruction		
6. Looks at page 4r	Looking for how to pass to level III instruction	
7. Operations manager asks him to apply sheets T204 and T205 to recover the auxiliary feedwater supply pumps	Deal with the AFS loss through application of recovery sheets	**Handling the auxiliary feedwater supply loss**
8. Takes sheets T 204 and T 205		

Figure 2: A Fragment of the Supervisor's Activity Reduced Account

In this extract, the first sequence is relative to the monitoring of level II operating instruction. The Supervisor is aware that the situation requires the application of the level III instruction, because one of the steam generator (SG) is radioactive. However, he can not take this level III instruction until the turbine operator has finished isolating the SG and tick off the state of this SG in a specific box of a crew reference document. At the same time, the Supervisor has noticed that, shortly before and thanks to its procedure that the AFS is no longer functioning; he must at the same time solve this specific problem (macro-sequence "handling the AFS loss") and integrate it in the rest of the operation (sequence "ensuring the request for cooling is appropriate with the AFS loss").

Furthermore, these operating documents are primarily designed to assist a 'single user' of procedures. A division of tasks is assigned to the operators and formal times of co-ordination between crew members are prescribed in the procedures. However, this pre-determined co-ordination is not sufficient: a specific job of articulation is implemented by monitoring the course of actions of the other crew members in order to make everyone's work easier. In this sense, operators must co-ordinate collaboration with the ongoing individual activity of others.

In other words, this important part of activity during which other things are carried out besides the strict application of procedures consists actually of building the context of their interpretation and supplying a framework, shared by all, for their application. This ability to put the prescribed tasks into context is founded on operators' know-how and experience.

Bringing to light the significant structures making up the activity of the operators who deal with scenarios enables to inform different situations such as :

1. the procedure makes up an assistance support for problem-solving.
2. the problems are treated independently from the application of instructions.
3. the procedures hinder the understanding of the situation.

The typology of these procedure-using situations and their detailed understanding allows for the enhancement of both their design and the training given to the operators.

CONCLUSION

This type of methodology aiming at fine comprehensive analyses and modelling of the operating activity on simulator is fundamental for explaining the quantitative results derived from other methodologies or for supporting the results derived from more intuitive analyses (e.g. : experts judgement).

The approach described in this paper must be completed by a comparison of these analyses of emergency situations or simulated accidents with analyses derived from perturbed or even incidental situations on units which will enable a better understanding of the whole organisation without being restricted to a reduced operating team (to 4 to 7 people) and a reduced simulation of the additional potential resources by the instructors.

REFERENCES

Bannon L, Robinson M. (1991). Questioning Representations, in *Proceedings of the European Conference on Computer-Supported Co-operative Work* , Amsterdam, The Netherlands.

Dien Y. (1993) Safety and Application of Procedures or How Do They Have to Use Operating Procedures in Nuclear Power Plants?; *HT-54/93-19A EDF*; Bad Homburg Workshop.

Filippi G., Theureau J. (1993) Analysing co-operative work in an Urban Traffic Control Room for the Design of a Co-ordination Support System, *Proceedings of the Third European Conference on Computer Supported Co-operative Work,* G. de Michelis, Simone C, Schmidt K. (Eds) 13-17 September. Milan .

Goodwin C., Goodwin M. (1991). Formulating planes : seeing as a situated activity in Y. Engestrom and D Middleton (Eds) *Cognition and Communication at work.*

OECD Halden Reactor Project, Halden, Norway.

Heath C., Luff P. (1991). Collaboration and Control : Crisis management and

Halbert B., P, Meyer P. (1995). *Summary of lessons learned at the OECD Halden Reactor Project for the Design and Evaluation of Human-Machine Systems,* multimedia technology in London Underground Lines Control Rooms *Computer Supported Co-operative Work - ECSCW ' 91,* **Vol 1,** n°s 1-2 .

Hutchins E. (1995)., *Cognition in the Wild.* MIT Press.

RÉSUMÉ

ANALYSE DE L'ACTIVITÉ DE TRAVAIL POUR L'ÉVALUATION DE PROCÉDURE DANS DES SALLES DE CONTRÔLE SIMULÉES : LE CAS DES OPÉRATIONS D'URGENCE DANS LES CENTRALES NUCLÉAIRES

Le recours aux simulateurs est essentiel pour étudier les facteurs humains dans les situations d'urgence des centrales nucléaires. Afin d'assurer un couplage efficace entre des dispositions techniques — en particulier des procédures — et les équipes opérationnelles des centrales dans les situations d'urgence, EDF a mis en place un « observatoire » : un système de surveillance qui recueille des données sur simulateur. Cette communication est centrée sur une méthodologie de recueil de données appelée ANACONDAS. Elle met en avant l'intérêt de l'analyse qualitative pour expliquer l'utilisation des procédures dans les opérations d'urgence sur simulateur.

Contextual analysis of the nuclear power plant operators' utilisation of a disturbance management system

Leena Norros & Maaria Nuutinen

VTT Automation, P.O.Box 1301 02044 VTT, Finland.
Leena.Norros@vtt.fi, Maaria.Nuutinen@vtt.fi

ABSTRACT

The scope of this study was to analyse if a new information tool for disturbance handling promotes control of the nuclear power plant process. A validation study was carried out on a full-scope training simulator. The results manifest the complexity of adopting new tools in professional practice. No significant general improvement in the adequacy of process control performance was found. However, the analysis of the working practices indicated that the tool may support the shift supervisor to organise the crew's performance. Moreover, the tool was felt to reduce stress in the situation. Diffuse task contexts revealed possible weaknesses in the tool to promote the shift supervisor's feel of the process and of the crew's performance.

Key words: process control, disturbance management, information aid, habits of action, contextual method.

INTRODUCTION

The typical tools in the modern process control industry are the automation and information systems. Essential to their use is the mediated character of the information they provide of the process (Zuboff 1988). The appearance of the process mainly in alpha-numerical form puts demands on intellectual abilities such as interpretation and judgement. Moreover, through its nature of being a human-made artefact an information system is also a model of the reality. Due to the huge amount of information available of the processes further conceptualisation is necessary in the design of human-machine interfaces. The model of the process is materialised during design through selecting and structuring information.

Because human beings do not only create tools but are themselves also shaped by these tools (Heinämaa & Tuomi 1989), it is necessary to reflect on the effects of the tools on the perception of the environment. This applies also to the NPP domain, where it is important to know how information tools shape the process operators' schemata of perception, and do the tools promote control of the process. Studies carried out for validation of different kinds of control room information systems aim to answer these questions. However, because it is difficult to find out the significance of such systems for perception of the process, these studies often restrict themselves to measurement of the effect of the system on some external performance criterion. This was also our starting point when faced with the task to study the usability of a particular information system that was aimed to aid NPP operators' management of disturbance situations. The system will here be called an information aid for disturbance management (IADM).

The purpose of the IADM system was to provide an overview of the plant state in a severe disturbance or accident situation and to aid the use of the emergency operating procedures. The IADM system consisted of two large conventional panels located in one side of the control room next to the normal operating panels. The information panels were well visible from the normal control positions of the crew all the time, and in a disturbance situation the shift supervisor was supposed to go over to the panels to get a closer look at the information provided by the system. One of the two panels presented information of the central process parameters of the plant in an illustrative way, i.e. the reactor and containment was depicted in a schematic way on the panel including the main components and the safety critical systems. The parameters relevant to the functioning of these systems were shown next to them through regular instruments. The other panel was an alarm system that provided information of the state of the critical safety functions of the process. It also indicated the access criteria to emergency operating procedures in case any of the critical functions were threatened. The system was an own design and it was retrofitted into the existing control room. It was also available in the full-scope training simulator.

THE RESEARCH METHODOLOGY

The study was carried out on a full-scope training simulator by a multi-disclipinary research team including also a simulator trainer and an expert of the power plant (Norros et al. 1997). The system was tested in four different severe accident situations which were to be realistic. The operators should have a possibility to affect the course of events in a positive way, but the disturbances should be complex and difficult enough from the point of view of diagnosis and operations. They should require co-operation among the crew and also contacts outside the control room.

Six NPP crews took part in the study. The crews consisted of four persons, a shift supervisor, and three operators. Each experimental session lasted two days during which, first, the purpose of the study was explained and discussed. Then two scenarios per day were run (1-2 h each). A final interview concerning the crew's conceptions of the information aid was also carried out.

In the study we used an experimental design according to which each of the six crews performed all four scenarios. Half of the scenarios were to be managed with and half without the extra information aid (with necessary permutations in the ordering of experimental conditions). The availability of the disturbance management system was the major experimental variable (available/not available) and the efficiency of process control the dependent variable.

During each run major parameters of the processes were registered. Video-recordings of the course of actions were gathered. Expert observations were also made of the course of actions according to a structured protocol. A debriefing interview with the crew was also carried out. After each run the research group went through the recordings for making necessary clarifications of the course of actions. The process experts prepared an evaluation of the adequacy of process control.

Based on this data and with the help of our methodology (Suchman 1989, Engeström 1987, Harré & Gillet 1994, Klemola & Norros 1997, Hukki & Norros 1998, Norros & Klemola in press) process control could be analysed comprehensively and in a context-dependent way. The analysis bases on a *systemic modelling of the operating contexts* (Hukki & Norros 1994, Holmberg et al. in press). First a description of the course of events was written. Thereafter, the process events were conceptualised from the point of view of the underlying safety critical functions of the NPP process and the long term boundary conditions for the usability of the plant. The critical functions and usability constraints set goals and task requirements for process control. These were made explicit with a further set of models. The first type of model is a functional flow-model of the task. In order to be able to define the concrete demands on the crew in the studied event, the specific circumstances of the situation, i.e. the situational possibilities to make inferences and to operate are described in relation to the task. These are provided in further models in which the task-related critical information of the event and its source are presented and the diagnostic and operational meaning of this information described. Likewise, the availability and usability of each alternative operating method is defined.

The operators' own accounts of their performance achieved through the debriefing interviews were essential for making the performance intelligible for the researchers. In the first phases of analysis a temporal description of the *operator-process interactions,* in the form of observations and operations was made, ordered under three broader categories: identification of loss of process stability, use of stabilisation measures, and identification of the cause of disturbance Based on this the adequacy of process control and the habits of action were evaluated with the help of the following criteria.

Adequacy of process control refers to the successfulness of the process control performance from the process point of view. Several aspect were considered: identification of the process state, stabilisation of the process, identification of the causes for the disturbance, use of instructions, summarising the situation with the crew and contacts outside. Scenario-specific operationalisations for these criteria were defined with the help of plant personnel who had substantial operating experience. In the evaluation of the adequacy of process control the process experts' ratings were used. For statistical analysis a sumvariable was formed of the first three items.

According to our methodology we also describe each crew's *habits of action*, which we define as the way of taking account of the constraints and possibilities of action expressed in the way of using situationally available resources. Habit of action is understood as an experientially acquired intentional and adaptive reaction towards the environment. Habits of action was considered from three aspects a) way of decision-making (global view of the situation, understanding of the character of the disturbance, taking account of situational constraints and action possibilities), b) way of co-operating (a shared interpretation of the situation and coherent team performance; and c) operator's personal way of coping with problem situations (reorienting in the problem situation, critical evaluation of own resources). Altogether 34 items were used as criteria for the assessment of the crews' habits of action. Also these criteria were operationalised with the help of the scenario-specific models of the operating context. The evaluation dimensions behind the items express to which extent the operators have in their diagnostic and operative decision-making urged towards a global interpretation of the particular disturbance, in their teamwork towards a shared interpretation and coherent co-operation, and in coping with problems towards a situationally adequate mobilisation of resources (Norros 1997). A more detailed description of the method is under preparation (Norros & Hukki under rewriting). Ratings were carried out by two researchers who through discussions attempted to achieve a consensus. For statistical analysis sumvariables were constructed based on the above three aspects of habit of action.

In the analysis of the results statistical tests were used with the intention to acquire an overview of the data and to sketch the results. With this background further questions were generated and qualitative inferences carried out for the conceptualisation of the organisation of the crews' actions.

RESULTS

The central question to be answered in the study was whether the information aid has an effect on the adequacy of the crews' process control performance.

For the first we checked if there were differences between the crews' overall performance level. The results show that there were no significant crew-dependent differences in the adequacy of performance. Then the main question concerning the effect of the information system on the adequacy of performance was studied. No statistically significant general effect was found.

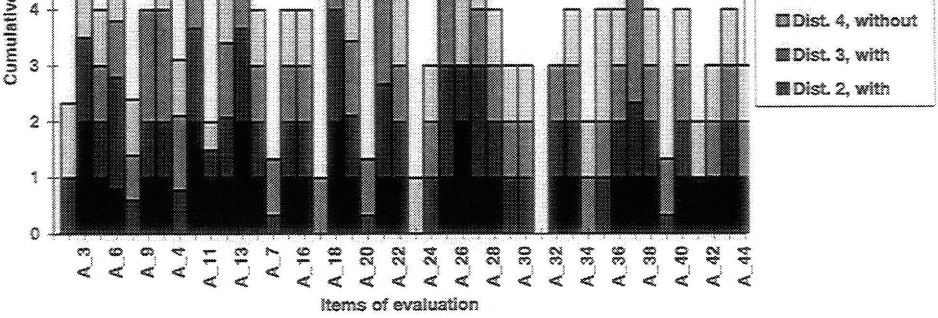


Figure 1. The habit of action profile of one shift in four disturbances.

We were, however, prepared to go further in the analysis. This required a restatement of the original question, which focused on the external relations between an input and output, between the availability of a new tool and the adequacy of performance. Instead, it could be asked how the operators use the system as a tool of an intentional action. When adopting this point of view it is natural to describe how the available resources are used, i.e. what kind of habits of action the operators have developed to cope with the process and are the advantages of the information aid dependent on the circumstances, here on the different disturbances.

On a global level the crews varied in regard to habits of action. This statistically assured result (in regard to way of decision-making $p<0.05$, in regard to way of co-operating $p<0.01$) indicates that habit of action was not very sensitive to situational differences, here to type of disturbance or information presentation. This fact can be demonstrated with an example of one shift's habits of action profiles of the four disturbance situations (figure 1). The ratings converge causing typical peaks and valleys in the profile.

The stability of habits of action is a reasonable result, because, by definition, habit of action is an experientially acquired intentional and adaptive reaction towards the environment, and, if appropriate, it takes situational factors well into account. This also implies that an appropriate habit of action should provide a good result in process control. A regression analysis indicated that habit of action explains 67% of the variance in the level of adequacy of process control. The effects of the way of decision making, $p<0.05$, and the shift supervisor's management practices, $p<0.01$, were statistically significant. Caution is necessary in interpreting the results because the anchors used in operationalising process control and habits of action criteria were partly overlapping, which caused internal dependency in the way of decision-making. This did not trouble the results regarding way of co-operating but there the evaluators themselves may tend to create dependency.

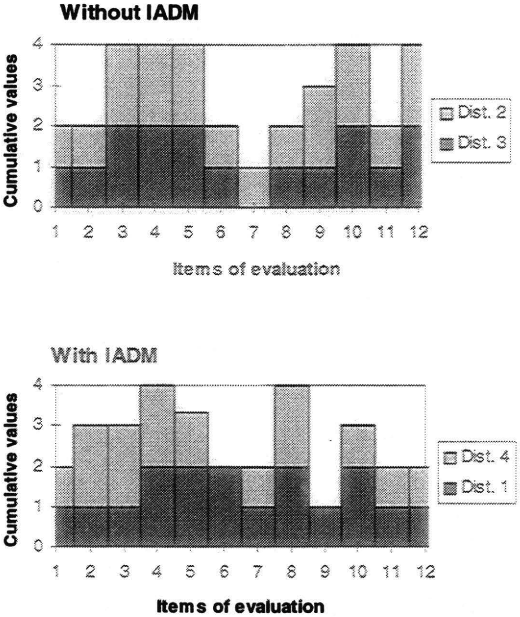


Figure 2. One shift supervisor's way of co-operating in two different disturbances without and with information aid (IADM).

Co-operation, in particular the shift supervisor's management practices seem to be significant in the control of disturbances. Therefore, it would be interesting to see if the type of information presentation possibly modifies these practices. Here we are not able to use statistics but rather scrutinise the phenomena in a qualitative way. We shall first take an example of one crew's results (figure 2).

In the upper part of the figure 2 we see the shift supervisor's way of co-operating in two disturbances that were managed without the information aid, and, in the lower part, with the aid. The figures show cumulative values of two evaluations (rating from 0-2) of each item. Thus, the value 2 expresses a sufficient level of co-operation, while a higher than that would indicate strengths and lower weaknesses in the co-operative practices. When comparing the two circumstances (without /with aid) we may notice that in the "with aid" conditions there is improvement in regard to items 2 and 8, while 4 remains high. These items are related to explicit organisation of the crews' actions. In regard to items 3 and 10, which express sensitivity towards and promoting the crew members' situation awareness, there is slight reduction when working with the IADM. The same is true in regard to items 5 and 12, which express sensitivity towards the global state of process and communicating reasoning based on it. The interviews with the crews put more weight on these observations. It was stated by he crews that the IADM helps in the operational structuring of the supervisor's own and the crew's actions. All shift supervisors felt that the availability of the IADM reduces stress in the disturbance situation. However, it was also stated that the IADM might draw the shift supervisor's attention from the on-going process events and crews' performance, particularly when using the alarm panel with access criteria for emergency operating procedures.

The conclusion of these results may be the hypothesis that the particular advantage of the information aid is in its ability to promote the shift supervisor's management actions and sense of control, expressed as reduction of stress in the situation, which both indirectly promote the disturbance handling. It is however possible that the slight loss of sensitivity of the supervisor towards the on-going situation may be disadvantageous. This possibility cannot be excluded by our results which showed tendentially that in three disturbances, characterised through a diffuse disturbance image, slow progress and a reduced availability of information, both the ratings of habit of action and the adequacy of process control were worse when working with the information aid than without. In the disturbance that was characterised by a clear disturbance image, that was evolving rapidly and in which there were no major deficiencies in the availability of process information the result was opposite, the information aid was related with better results.

DISCUSSION

As noted in the introduction, the fact that human beings not only create tools but are also shaped by these tools, makes it necessary to reflect on the effects of the tools on the perception of the environment. Validation studies aim to answer these questions but they often restrict themselves to measurement of the effect of the tool on some external performance criterion. We took a step forward and attempted to analyse changes in the dynamical structure of action when adopting new tools. This point of view was considered more informative for the prediction of the future appropriateness of the tools, because in complicated work an attempt to find straight forward changes in some simple performance criterion is not very realistic. For the first it is difficult even to experts to agree what such criteria could be. Secondly, expert level actors who are usually involved in validation test are able to compensate unoptimal conditions. Thirdly, only in trivial cases the possible benefits of the new equipment are so evident that experts are able to utilise them on first trial, or after a short introduction. The description of the habit of actions through a detailed analysis of the courses of action in the task situations, was assumed to reveal the dynamics of actions. The habits of action were found to be rather stable across the experimental conditions but, nevertheless, also signs for changes related to certain characteristics of the information aid were found out. Further analysis of these indications would be the strategy of making use of the possibilities of the new tool and of preventing its eventual disadvantages. Thus, reflection and development of the habits of action should be considered the method of implementing the new tools in practice.

When evaluating the validity of new tools, it is generally found important to use a variety of test situations, which may manifest unanticipated demands on the tools. Yet, it seems less frequent to analyse the differences in the test situations, which means that the variance of the results remains unexplained (Bove & Andersen 1999). In this study we made effort to describe the task features in a systemic way as contexts of action. This was assumed to allow inferences regarding relevant general features of the applicability of the tools (Hukki on going work, Holmberg et al. in press). We think however, that the advantages of the analysis of the different task contexts, in which the characteristics of tools may be revealed, are realisable only through the analysis of actual habits of action. This is due to the fact that while a particular feature of a tool may become manifest in a situation, this information becomes meaningful only when related to the functionally definable adaptivity demands on behaviour, which are the evaluation categories for the habits of action (Norros & Klemola in press, Norros & Nuutinen 1999). In the present case, the insight that it is an essential characteristic to an appropriate habit of action to strive towards a global interpretation of the particular situation through taking account of its specific features, helps to see that the inability of the information aid to support this goal may be a significant deficiency in task situations that are "untypical". Through this insight it was also possible to concretise the meaning of the claim that information tools are conceptual generalisation of the process. If the particular, situation-specific behaviour of the process corresponds well with the system, it supports the interpretation and control of the situation well. If the correspondence is weaker, the support may in this respect be smaller. The difference between the epistemic nature of tools and that of the particular situations of their application was recently emphasised also by Beguin et al. (1999). This fact should be taken into account as a general condition of the applicability of information aids, and made explicit in training because it promotes a realistic attitude towards the tools.

REFERENCES

Beguin, P., Kazmierczak, M., Cottura, R., Leininger, J. & Vicot, P. (1999). Design of an alarm and risk managment in chemistry. Some lessons learned from interdisclipinary research. Pre-Proceedings of Human Error, Safety, and Systems Development, Liege, June 7-8, 1999.

Bove, T. & Andersen, H.B. (1999). The effect of advisory system on pilots go/no-go decision during take off. Pre-Proceedings of Human Error, Safety, and Systems Development, Liege, June 7-8, 1999.

Engeström, Y. (1987). Learning by Expanding. An activity-theoretical approach to developmental research. Helsinki, Orienta-Konsultit.

Harré, R. & Gillet, G. (1994). The Discursive Mind. Thousand Oaks, Sage Publications.

Heinämaa, S & Tuomi, I. (1989). Thoughts Creating Machines. Dreams and Nightmares of Artificial Intelligence. Juva: Werner Söderström Oy (in Finnish).

Holmberg, J., Hukki, K., Norros,L., Pulkkinen, P. & Pyy, P. (in press). An integrated approach to human reliability analysis - decision analytic dynamic reliability model. *Reliability Engineering and Systems Safety*, 65, 239-250.

Hukki, K. & Norros, L. (1994). A method for analysis of nuclear power plant operators' decision making in simulated disturbance situations. In: Norros, L. (ed) XIII European Annual Conference on Human Decision Making and Manual Control. VTT Symposium 146. Espoo: Technical Research Centre of Finland, 200 -217).

Hukki, K. & Norros, L. (1998). Subject-centred and systemic conceptualisation as a tool of simulator training. *Le Travail Humain*, 4,313-331.

Klemola, U-M. & Norros, L. (1997). Analysis of the clinical behaviour of anaesthetists: Recognition of uncertainty as a basis for new practice. *Medical Education*, 31, 449-456.

Norros, L. (1997). Human factors in NPP operations. In: Vanttola, T. & Puska, E-K. (eds) The Finnish research programme on Reactor Safety, Interim report 1995-1997. Research Notes 1856, Espoo: Technical Research Center of Finland, 146-159.

Norros, L. Hukki, K. (under rewriting). Dynamics of process operators' decision-making in a disturbance situation. A contextual analysis. *International Journal of Cognitive Ergonomics*.

Norros, L. & Klemola U-M. (in press). Methodological considerations in analysing anaesthetists' habits of action in clinical situations. *Ergonomics*.

Norros, L. & Nuutinen, N. (1999). Development of an approach to analysis of Air Traffic Controllers' working practices. Pre-Proceedings of Human Error, Safety, and Systems Development, Liege, June 7-8, 1999.

Norros, L., Holmberg, J., Hukki, K. & Nuutinen, M. (1997). Validation of NPP control room information aid. VTT Automation. Unpublished report.

Suchman, L.A. (1987). Plans and Situated Actions. The problem of human-machine communication. Cambridge: Cambridge University Press.

Zuboff, S. (1988). In the Age of the Smart Machine.Oxford: Heinemann Professional Publishing.

RÉSUMÉ

ANALYSE CONTEXTUELLE DE L'UTILISATION D'UN SYSTÈME DE GESTION D'INCIDENT PAR DES OPÉRATEURS DE CENTRALE NUCLÉAIRE

MOTS CLÉS : Contrôle de processus, Gestion des incidents, Assistance à l'information, Habitudes d'action, Méthode contextuelle.

User acceptance evaluation in real operational conditions: the case of FARAWAY data link communication system

Patrizia Marti, Marco Palmonari
Multimedia Communication Laboratory
University of Siena
{marti/palm@media.unisi.it}

Guglielmo Casale
Alenia Radar System Division
{gcasale@lti.alenia.it}

ABSTRACT

This paper describes the user acceptance evaluation of an innovative communication system in the air traffic control domain. The evaluation of the system, namely a bi-directional ground-air data link communication system, was carried out in real work environment in parallel with the development and the assessment of technical aspects; for these reasons, the main aim of the user acceptance was to estimate the impact of the new data link technology applications on the controllers' and pilots' activity.

Keywords: air traffic control, data-link, user acceptance evaluation, distributed cognition

INTRODUCTION

Air Traffic Control (ATC) is aimed at assuring a safe and expeditious flow of air traffic. Indeed, the controllers' main objective is to maintain appropriate separations between aircraft; the entity of the horizontal minima separation varies according to the different sectors: from 5 to 10 miles for en-route sectors (the upper sectors remote form the airports) and 3 miles for terminal area (the area close to the airport).

At present, air traffic surveillance is performed by means of two main artefacts, the radar which displays the current aircraft behaviour within a given sector and flight strips, that represent aircraft past and future behaviour in terms of a flight plan. Controllers use verbal communication (through a radio frequency) to communicate with pilots. In particular, radio frequency is used by controllers for two main purposes, i) to act upon the traffic in order to prevent or solve conflicts and ii) to obtain information that are not provided by the radar (aircraft air speed, lateral deviation, rate of climb or descent). Similarly, radio frequency is used by pilots with two different aims, i) to ask for a flight plan variation (e.g., short route to the airport of destination) and ii) to obtain information that are not provided on board (e.g., weather data).

Radio frequency congestion is considered the main bottleneck of current ATC systems because it heavily constraints the amount of manageable aircraft. For this reason, radio frequency does not seem to respond to the increasing traffic demand.

Data link (DL), a technology of digital broadcast communication, seems to be able to enhance both ATC system capacity and safety. In particular, DL systems are aimed at supplementing the voice communication by providing an additional communication means that reduces congestion on voice channel. DL systems can have different relevant applications in ATC domain and different uses of data link can be envisioned (FAA, 1988, 1989).

First, data-link can be used by controllers to act upon the traffic; the rational underlying this application of data-link technology is to significantly reduce the use of the radio channel for routine communication. Controller-pilot data link communication can include a set of clearance/information/request messages corresponding to voice phraseology currently employed for voice communication. The main problem related with the introduction of DL is that the time needed to formulate and send the message using DL is longer than the time needed to issue the same instruction by voice. For this reason, DL is not completely appropriate to substitute voice communication in time critical situations, that is situations (such as short term conflict resolution) requiring a rapid controllers-pilots communication. Second, data-link technology can be used for automatic air-ground data exchange; for instance, data-link can be used to provide controllers with relevant data on traffic, as aircraft air speed, that are not provided by radar systems. Similarly, weather advisories can be automatically sent from ground systems to aircraft. As described in the present paper, this second application of DL technology can have a positive impact on controllers and pilots activity and can contribute to the enhancement of ATC system capacity and safety.

AIM OF THE STUDY AND THEORETICAL APPROACH

The paper describes the user acceptance evaluation of an innovative bi-directional ground-air data link communication system designed to provide an automatic air-ground data exchange.

The evaluation, which was carried out in real work in parallel with the development and the assessments of the technical aspects, aimed at providing a first assessment of the impact of the new data link technology applications on the controllers' and pilots' activity.

The theoretical approach which informed our analysis is Distributed Cognition (Hutchins, 1995, 1996). According to this approach, a socio-technical system such as ATC can be analysed as a cognitive system composed of groups of individuals (pilots and controllers) interacting with each other through a set of artefacts (radar, radio frequency, strips, radio frequency) over a period of time. This cognitive system provides the information resources needed to inform the control activity. These information resources are represented internally (within the individual mind of the different actors involved in the process) or externally (on different artefacts or media environment). Resources which are externally represented take different expressions: strips, verbal communications and radar screen are examples of different media representing information on the air traffic (Fields, Wright, Marti, and Palmonari, 1998).

The central tenet of Distributed Cognition approach is that any process within a socio-technical system can be analysed in terms propagation of representations across different media. According to this view, process' bottlenecks and breakdowns depend on the properties of the media which, in turn, determinate what kind of information the different actors involved in the process can access and elaborate to perform the task.

THE CRITICAL ISSUE: BOTTLENECK IN INFORMATION MANAGEMENT

The situation

In order to better clarify the critical issue that is discussed in the paper, a real scenario describing the activity carried out in the Ciampino operational control room (Rome) is presented (see table 1). In particular, the scenario describes the activity of two controllers in the approach sector to the Fiumicino airport (Rome), namely:

- the planner controller, who mainly plans the traffic flow within the sector;
- a tactical controller, who monitors the traffic within the sector and guides pilots issuing clearances and guidance through radio frequency according to the plan elaborated by the planner controller.

Situation's description	Radio frequency communications	Controller's comments
The tactical controller is coordinating with the planner controller and does not pay attention to the Swissair pilot's call (on the radio frequency); the pilot, after few seconds, calls again and asks for confirmation of the approaching procedure; finally, the controller clears the pilot to follow the procedure.	**Pilot:** "Hello..this is 3602.." **Pilot (after few seconds) :** "..Rome 3602" **Controller:** " 3602 Rome ?" **Pilot:** "...Rome 3602..confirm the approaching procedure ? **Controller:** "Yes, you are number 2..."	"The Swiss pilot was asking me if he was cleared to approach...I probably was co-ordinating (with the tactical controller) and I didn't hear his call..probably he will call me again shortly." " Yes, he is calling.."

Table 1 : "Controller Multitasking".

Comments

The scenario describes how the properties of the communication media affects controllers performance: sometimes controllers cannot pay attention to all the information sources which are relevant at a given time, in particular when the elaboration of the information coming from the different sources (pilots and controllers) involves the same sensory modality (the auditory modality).

In general, the verbal communication among controllers and between tactical controller and pilots has two drawbacks:

- the bandwidth limitation of the communication medium (just one pilot can occupy the radio frequency and communicate with the controller at once);
- the limitations of the controllers' elaboration skills (as described in the scenario, the tactical controller cannot pay attention to the radio frequency and to the planner controller at the same).

THE FARAWAY PROJECT

FARAWAY was a project funded by the European Commission in 1997 within the Telematics Application Programme. Its objective was to develop a bi-directional ground/air data link system based on ADS data. The idea is that during the flight such a system automatically provides pilots and controllers with different kind of data (e.g. air speed, climb/descent rate, meteorological data) with the double aim of reducing the need of direct verbal communications via radio frequency and improving the data accuracy.

Whilst the system was not aimed at substituting verbal communication, it moreover sought to distribute the information exchanged during the flight across different media (on board display, data integration on the radar screen on ground).

The project ended in 1998 and demonstrated the maturity of the proposed technological solution and the first results about the impact that such technology could have on the human activity.

The project was successful so that the European Commission funded in 1998 a follow up project named FARAWAY 2.

EVALUATION PROCESS

FARAWAY was mainly a "technology driven" project. As stated above its main objective was to demonstrate the maturity and the effectiveness of the developed technology. For this reason, the system was tested in real operational conditions. Alitalia MD80 aircraft were equipped with the system on board, and about 20 flights were tested. Notwithstanding the technological feature of the project, the evaluation of the impact of such a system on the human activity acquired an increasing interest during the trials. For this reason a considerable effort was devoted to set up a user acceptance evaluation in accordance to the main technical evaluation requirement: *to test the system in real operational conditions.*

The evaluation was planned according to the constraints of the project; indeed, for safety reasons, the pilots and controllers involved in the evaluation could not really use the system during the operational trials; in addiction, the user interface was not the focus of the project in that it was developed just for giving an "indication of use".

User evaluation process was articulated in three main phases.

1. *Data collection and analysis*, during which we performed ethnographic field studies, interviews, user analysis, task analysis,
2. *Operational trial*, during which we video recorded the activity in the cockpit and on ground (the Ciampino control room), and performed an "opportunity evaluation" with pilots and controllers during and after the flights;
3. *Synthesis and envisioning*, during which we organised participatory workshops with the users, showing and commenting the video recorded in phase 1.

The main steps of the user acceptance evaluation process are briefly discussed in the following.

Users and stakeholders analysis

The target user of the system were identified in two categories: end-users and stakeholders. Our end-users are pilots and controllers, whilst our stakeholders are the representatives of Alitalia (the Italian airline) and ENAV (the Italian organisation for air traffic management) managing directions. In the preliminary phase of the users assessment evaluation, non-structured and semi-structured interviews with end-users were carried out to get a general overview on the usability and effectiveness problems of the new system and to collect data about pilots' and controllers' activity with the aim of defining use scenarios. These data were also input to activity analysis.

Activity analysis

Activity analysis was carried out on the basis of observations, interviews, video and audio recordings. The analysis, both qualitative and quantitative, was mainly devoted to the information exchanged through the VHF. In fact radio frequency allows tactical to guide aircraft and to get information about:

1. The traffic. Some data like air speed are not displayed on the radar screen, and would require a mental overload to be obtained otherwise;
2. Pilot's intentions and feelings. Again the information are not provided by any other medium in the environment.

On the other hand, the radio communication is also used by the pilot to request changes of route and to get information about the procedures to be followed.

Qualitative analysis

Activity analysis allowed to understand the importance of information redundancy in ATC. Indeed, the possibility for all the pilots to access through the radio frequency the information concerning the surrounding traffic and the controller's activity is a crucial property for the ATC system because it enhances safety by supporting the co-operation among controllers and pilots. This property is extremely relevant when introducing a new communication technology as the data link which, designed to decrease the radio frequency loading tends to reduce the amount of information about the state of the traffic accessible by the pilots; to preserve safety standards, the new technology should somehow provide pilots with information concerning the surrounding traffic.

Quantitative analysis

A quantitative analysis of the radio frequency communication among controllers and pilots in an en-route sector was also carried out. The analysis, based on the record of few hours of communication, was aimed at identifying which categories of communication could be eliminated by providing both pilots and controllers with DL data. In the following we present the frequencies of occurrence of communication categories relevant for the use of DL data, the total amount of communication exchanged being 163.

Pilots

- requests for change of flight level (example: "Alitalia 1898..can climb at 310 ?"): 5;

of which refused by the controller: 3;

- requests for change of route (request for a shorter route, example: "Alitalia 1760 request after Latina direct to Bolsena"): 7;

of which refused by the controller: 3;

- requests for information: none

Tactical Controller

- Proposition/guidance for a direct without the explicit request of the pilots (example: "4007 if you want a shorter route to Tarquinia for us is ok") : 6
- requests for information (speed, example: "How much speed now?"): 2
- requests for information (level the a/c has to reach, example: "Alitalia 1898..what is your level?"): 3

This preliminary quantitative analysis allowed to get a first account on the possible impact of DL on controllers-pilots communication. In particular, among the categories presented above, at least two could be eliminated thanks to the use of DL data:

- the pilots request for a change of flight plan (both altitude and route) in conditions of high traffic (the requests that the controllers refuse);
- the controllers requests of information concerning the a/c (e.g., speed).

Use scenarios

During the analysis phase we selected together with the users a number of current work situations and developed use scenarios for potential future system support (the FARAWAY applications). Each scenario was intended to detail a usage situation and to document step-by-step actions performed by the users (Carrol, 1993). The first set of scenarios included typical sequences of actions and events that occur during different phases of the flight. Each of these phases had its own drawbacks that were stressed by the users involved in the definition of use scenarios. The use scenarios had the objectives to get a common understanding of the activity of controllers and pilots, and to set the stage for the first user focus group. Both current work situations and future use scenarios were further revised for the second focus group.

Participatory workshops

Due to the need to maintain safety standards during the evaluation phase it was impossible to set up a real or even a simulated setting involving all the actors. When it is not possible to carry out real use tests in the early phase of the system's specification, the focus group is one of the most suitable method to assess the impact that a new system would have on the activity of its different final users.

Two focus groups were carried out with the aim to collect evaluative and envisioning comments about the system from the different end users involved in the its evaluation. The group was composed of pilots and controllers who tested the system, Alenia system developers, Alitalia representatives and two facilitators.

Focus groups played a crucial role in the evaluation process in that they allowed to get detailed and accurate comments from the pilots and the controllers on the impact of the system on their activity.

FINDINGS AND COMMENTS

Three main issues were raised during the user acceptance evaluation. For the sake of clarity the issues are presented separately even if they are tightly related each others.

Impact on the pilots' activity

Currently pilots do built up a representation of the surrounding traffic using the information they can access through the radio frequency; nevertheless, this activity is very demanding in terms of attention and memorial resources and can interfere with the other activities the pilots have to carry out. The operational trails carried out within FARAWAY showed that the visualisation of the surrounding traffic on a display should both decrease the pilots' workload and increase the quality of their representation of the traffic.

Impact on the controllers' activity

Data link system using ADS data can provide the controllers with information about the state of the traffic which is not currently directly available (i. e. , the aircraft's descent rate). The analysis demonstrates that ADS data could support the controllers' activity in two complementary ways:

- by enhancing the controllers' awareness of the traffic current state and evolution;
- by decreasing the controllers workload; indeed, currently the controllers either mentally calculate the data they need or ask the pilots for the information.

Impact on the communication controllers-pilots

The analysis carried out within FARAWAY shows that DL data, by enhancing the controllers-pilots traffic awareness and co-ordination, can help in decreasing the radio frequency load in three ways:

- reducing communication due to misunderstandings between controllers and pilots;
- reducing pilots' request of changes in the flight plan when the traffic conditions do not allow it;
- supporting the pilots' anticipation of the traffic evolution and "self-spacing" thus reducing "routine" controllers' interventions on the traffic.

FINAL REMARKS

The user acceptance evaluation process carried out within FARAWAY project provides preliminary evidence supporting the assumption that the use of DL technology for automatic air-ground data exchange can increase ATC capacity and safety. The idea underlying this application of DL technology is that the distribution of the information currently exchanged by means of radio frequency across different media (on board display, data integration on the radar screen on ground) can at the same time contribute in decreasing the congestion of radio frequently and enhancing both controllers and pilots awareness of the traffic.

REFERENCES

Carroll, J.M. (1993). *Scenario-based Design.* New York: Wiley & Sons.

Clegg, C.W., Warr, P.B., Green, T.R.G., Monk, A., Kemp, N., Allison, G. and Lansdale, M. (1988). *People and computers - how to evaluate your company's new technology.* Ellis Horwood: Chichster.

Dumas, J.S. & Redish, J.C. (1993). *A practical guide to usability testing.* Norwood: Ablex Publishing Corporation.

EUROCONTROL. (1995). "A Human-Machine Interface Reference System for En-Route Air Traffic Control". *EEC Report n. 292: EEC Task AT66*, EatCHIP Task DPS.ETI.ST07. (http://www.eurocontrol.be/projects/eatchip/hmi/).

Fields, R., Wright, P., Marti, P., and Palmonari, M. (1998). Air traffic control as a distributed cognitive system: a case study of external representations. T. R. G. Green, L. Bannon, C. P. Warren and J. Buckley. (Eds). *ECCE98 Proceedings of the Ninth European Conference on Cognitive Ergonomics.* Limerick, Ireland.

FAA. (1989). Controller Evaluation of Initial Data Link Air Traffic Control Services: Mini Study 2 Volume II, *report n. DOT/FAA/CT-89/14, II.* (http://www.tc.faa.gov/act300/act350/act350.html).

FAA. (1988). Controller Evaluation of Initial Data Link Air Traffic Control Services: Mini Study 1 Volume 1, *report n. DOT/FAA/CT-88/14, 1.* (http://www.tc.faa.gov/act300/act350/act350.html).

Hopkin, V.D. (1995). *Human Factors in Air Traffic Control.* London: Taylor & Francis,.

Hutchins, E.. (1995). "How a Cockpit Remembers its Speeds". *Cognitive Science*, 19, 265-288.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: the MIT Press.

Kyng, M. (1994). "Scandinavian Design: Users in Product Development". *Proceedings of CHI 1994.* (pp..3-9). Boston, MA.

Norman, D. and Draper, S.W. (eds). (1986). *User Centred System design: New perspectives on Human-Computer Interaction.* Hillsdale, NJ: Lawrence Erlbaum.

RÉSUMÉ

ÉVALUATION DE L'ACCEPTATION PAR L'UTILISATEUR EN CONDITIONS OPÉRATIONNELLES RÉELLES : LE CAS DU SYSTÈME DE COMMUNICATION PAR LIAISON DE DONNÉES *FARAWAY*

Cette communication décrit l' évaluation de l'acceptation par l'utilisateur d'un système innovant de communication dans le contrôle du trafic aérien. L'évaluation du système, à savoir un système air-sol de communication bi-directionnelle par liaison de données, a été réalisée dans un environnement réel de travail, en parallèle au développement et à l'évaluation des aspects techniques. Pour ces raisons, le principal objectif de cette étude a consisté à examiner l'impact des nouvelles technologies de liaison de données sur les activités des contrôleurs et des pilotes.

MOTS CLÉS : Contrôle de trafic aérien, Liaison de données, Évaluation de l'acceptation par l'utilisateur, Cognition distribuée.

Managing Interactions Between Car Drivers : an Essential Dimension of Reliable Driving

Farida Saad°, Christophe Mundutéguy* & Françoise Darses*

°Laboratoire de Psychologie de la conduite, INRETS, Arcueil. saad@inrets.fr
*Laboratoire d'Ergonomie, CNAM, Paris. munduteg@cnam.fr; darses@cnam.fr

ABSTRACT

Whereas a lot of research has been devoted to the interactions between drivers and the road infrastructure, how drivers manage their interactions with other drivers have not, to our knowledge, been studied very extensively. Managing interactions nevertheless plays an important role in safe driving in most road situations. The purpose of our research works is to provide a deeper analysis of the means for managing interactions between drivers in car-following situation and to produce a preliminary model. It should eventually contribute to identifying means of communication that could facilitate the management of these interactions.

Keywords: Driving car, critical car-following, managing interactions, action prediction, uncertainty.

INTRODUCTION

Whereas a lot of research has been devoted to the interactions between drivers and the road infrastructure (cf. for example, the work on the mental categorisation of road sites – Mazet, 1991; Dubois et al, 1993), the means for managing interactions between drivers have not, to our knowledge, been studied very extensively. Managing interactions nevertheless plays an important role in safe driving in most road situations. A certain amount of research has highlighted the difficulties encountered by drivers in managing these interactions, and these difficulties explain the occurrence of certain accidents (Malaterre, 1990; Van Elslande et al, 1997) and traffic conflicts (Monseur & Malaterre, 1969; Risser, 1985). Different elements at the root of these malfunctions have been identified, such as the application of contradictory systems of rules (formal or informal) by the different participants in a situation, the lack of communication between users, or misunderstanding another person's behaviour or intentions.

This last factor seems to be at the heart of the management of interactions between drivers. In-depth analyses of activity in specific driving situations, such as crossing intersections (Saad et al, 1990) or car following on motorways (Saad, 1996), confirm that the way a driver behaves closely depends on his interpretation of other users' behaviour and his predictions of their intentions. These analyses also show that some regulating actions taken by drivers (reducing headway in car-following situations or maintaining speed in the presence of another user at an intersection, for example) are designed to communicate their intentions to other users and/or influence their behaviour.

The purpose of our research works is to provide a deeper analysis of the means for managing interactions between drivers and to produce a preliminary model. It should eventually contribute to identifying means of communication that could facilitate the management of these interactions.

DRIVING SITUATION STUDIED: CAR FOLLOWING ON AN URBAN MOTORWAY

A car-following situation is generally considered to begin once a driver can no longer drive at the speed he would like to because of the presence of other users in his lane (Leutzbach, 1974). Depending on the number of vehicles on the road and the traffic lanes available, this constraint may be more or less great and more or less long-lasting. Different degrees of density have been defined to characterise the state of the traffic (Carré et al., 1984), which range from freely circulating traffic to occasional constraints to heavy congestion. At each density level, there may be sharp and rapid local fluctuations in traffic flows.

Car-following situations call for the driver to adapt to the constraints and variations of the traffic, and in particular to detect "critical" variations, i.e. those which require a regulating action to restore adequate safety margins or prevent a collision. In this kind of traffic, where drivers' actions are closely interdependent, the degree to which a driver adapts his behaviour depends on the safety margins he adopts and the way he controls his interaction with other users. The action taken by another driver may more or less rapidly transform the variables characterising the driving situation and interfere with the tasks the driver is performing or planning to perform. Hence, to give some examples, when another user moves into his lane, the driver may have to rapidly readjust his safety margins; or if a driver ahead of him or behind him pulls out, that may deprive him of an opportunity he may have had to change lane himself; or if traffic slows down suddenly in the lane into which the driver intended to move, that may reduce the space that was originally available and lead him to defer his change of lane.

Managing interactions with other users efficiently calls for the driver to understand their driving and anticipate their intentions. Understanding and anticipating intentions depend on the information that the other drivers communicate either explicitly, through the use of formal signals (such as their indicator), or implicitly, through their behaviour (speed, positioning on the road, reduction of headways in car-following situations, and so on). They also depend on the driver's body of knowledge, which structure his expectations and enable him to formulate hypotheses about the adjustments that other users may force him to make in his driving.
How does this wealth of information at the driver's disposal come together to form significant elements for interpreting the intentions of other drivers? How, in a shared space, does the driver adjust his behaviour to take account of the actions of other drivers?

We offer an initial approach to these questions through two complementary studies. The first, carried out in a real driving situation, made it possible to identify the spectrum of variables that drivers take into account in managing their interactions with other users and determining their safety margins in car-following situations The second study took the interactive dimension of driving further and was aimed at understanding the mechanisms through which a driver recognises the intentions of other users and the information they communicate (whether directed at him or not). It was backed up by a laboratory experiment in which subjects were shown videotaped motorway scenarios and commented on the significant elements that enabled them to infer what action another driver would take.

MANAGING INTERACTIONS WITH OTHERS IN A REAL DRIVING SITUATION

The first study (Saad, 1997) was an in-depth investigation into driver behaviour in car-following situations, designed to identify the difficulties specific to this type of situation and the kind of assistance that could help drivers adapt to it. It aimed to:
- define more precisely the situational demands (infrastructure and traffic related) to which drivers have to adapt when driving in car-following situations;
- analyse how drivers take these situational demands into account and how they organise, perform and control the different tasks required.A specific objective was to identify the information and knowledge on which drivers based their categorisation of the situations they encountered and the strategies they applied in managing those situations (motives, criteria and decision thresholds for the type of regulating action taken).

Methodology

The method used combined two investigative techniques: on-board observation of driver behaviour during a journey on a motorway, and subsequent verbal reports and interviews. Each driver drove along the same predetermined route, for which a preliminary analysis of situational demands had been made. The route chosen for the study was a stretch of urban motorway, covering a distance of about 20 kms. The drivers used a vehicle equipped for video-recording the journey and for collecting data on a number of indicators, such as time (to date the event observed), distance (to situate it along the route), driver's speed, use of brake and so on.

Immediately after completing the journey, each driver was interviewed (yielding verbal reports on the video recording of the journey and semi-structured interviews). Two groups of drivers participated in the study - one of three experienced drivers and one of three novice drivers.

Main Findings

Analysis of driver behaviour : involvement in critical car-following episodes

The analysis of driver behaviour indicated that the safety margins observed in car-following frequently varied, very often reaching low values (≤ 1.5 s). A number of "critical car-following episodes" were identified : 33 of which involved experienced drivers and 40 novice drivers. An examination of these episodes enabled us to characterise both drivers' behaviour and the environmental conditions likely to influence their activity (level of traffic constraints and infrastructure characteristics, immediate interactions with other drivers).

The instability of critical car-following episodes

Most of the critical episodes observed were "unstable" ones, i.e. they were associated with a lane change manoeuvre performed either by the driver himself and/or by another road user. When traffic density was high, the critical episodes were long and sometimes associated with drivers' braking actions (this is mainly the case for novice drivers). The instability was observed particularly frequently along some stretches of motorway, transition zones that impose temporal and structural constraints on the tasks to be performed (lane-changing manoeuvres to follow an itinerary or to join another section of motorway). It was also in these zones, or in their immediate vicinity, that critical variations caused by other drivers pulling into the driver's lane were particularly frequent. Finally, it should be noted that several incidents or conflicts between the driver and other road users occurred in these zones. These observations confirm that areas of co-activity, where drivers have to co-ordinate theexecution of their own tasks with those of others, may represent "reliability blackspots" (Faverge, 1970).

The interactive dimension of driving in car-following situations

Our analysis also emphasised the interactive dimension of driving in such traffic conditions. Although the critical margins observed during the journeys very often resulted from the driver's own decisions (closing up on another driver before overtaking, for example), they were also partly imposed by the actions of other users. This is particularly evident when another user cuts in on the driver's lane. Such manoeuvres often result in rapid and significant reductions in the driver's margins. The integration of the resulting variation happens more or less quickly depending on traffic conditions. In heavy traffic, temporising strategies were particularly necessary, for if the driver wished to modify the margin with the preceding vehicle, he had to ensure that this correction was compatible with the distance of the vehicle following him. If pressure from the rear was too great, such variations could only be integrated progressively, which requires a certain time.

One-third of the critical sequences observed were directly associated with such cutting-in manoeuvres, and novice drivers were more frequently exposed to them than experienced drivers. Several elements may explain the frequency of such events for novice drivers : they travelled mostly in the middle and right-hand lanes and were thus more likely to be subjected to frequent lane changes on the part of other users (particularly in directional zones and traffic entry zones); when involved in car-following, they generally drove at lower speeds than that of the traffic flow; finally, in some situations (when they have to select their lane in directional zones or when merging on another stretch of the motorway), their uncertainty regarding future events or the choices to be made resulted in their leaving longer margins, thus providing other drivers with an opportunity to join their lane. This result suggests that novice drivers' involvement in critical episodes would seem to be due more to their difficulties in managing the complex interactions involved in car-following situations than to deliberate risk-taking. Experienced drivers seem more able to manage (and to avoid) such critical interactions, mainly because they are familiar with the conditions in which they are likely to occur. Some "anticipated" lane changes are indicative of this, for example, when a driver anticipates another vehicle joining his lane before the other driver has made his intention known by indicating, or when he changes lane on nearing a traffic entry zone before any other road user is actually visible.

Analysis of drivers' verbal reports : drivers' criteria for controlling safety margins

An analysis of drivers' verbal reports complemented the analysis of observed behaviour. It enabled us to identify the range of variables that drivers take into account in managing their interactions with other users and, in particular, their safety margins in car-following situations. They include: the characteristics of the infrastructure; the overall speed, density and stability of the traffic flow; the vicinity, extent and duration of safety margin variations; and the nature of the immediate interactions with (and between) other drivers. The regulating actions taken (or not taken) by the drivers depended, not only on the value of the margin with the preceding vehicle, but on the global traffic situation and its dynamics, as well as their intentions and priorities in the driving context. They also depends on other's behaviour or intentions (as observed or expected) and are sometimes aimed at influencing that behavior.

The results also suggest that, when managing their interactions with other road users, drivers (and especially experienced drivers) draw on a number of "reference situations" and base their decisions on a representation of the "typical behaviour" of other road users in those situations. If they detect any deviation from that norm, they intensify their monitoring of the situation and/or take some form of anticipatory regulating action (changing lane, reducing speed). One of the main problems from a safety standpoint, as is highlighted by the analysis of conflicts between road users, is to identify what information is required by drivers if they are to realise that the situational model on which they base their behaviour is no longer appropriate and needs to be updated.

INFERRING OTHERS' INTENTIONS IN A SIMULATED SITUATION

The purpose of this second study was to explore the way in which drivers make inferences about the intended actions of other drivers. The idea was to isolate the elements that one assembles in producing a representation of the behaviour of drivers that are, or are likely to be, interacting with oneself.

Methodology

This study did not deal with the activity of driving itself: 10 subjects were shown the videos recorded during the round trips made in real driving situation (see the first study above). The subjects were experienced drivers and did not know the route. The projection was interrupted when specific episodes of interaction occurred (marked in advance). Altogether, the subjects were shown 155 interaction sequences. At the end of each one, they were asked to comment on the intentions of the other drivers. They then had to describe what action they would have taken in that type of driving situation. These comments were recorded and transcribed (for further details, cf. Mundutéguy, 1999).

The first results presented here come from the analysis of these verbal statements. The objective of this continuing work is to construct a model of how drivers infer other users' intentions.

Main Findings

Intention is elaborated by combining categorisation and circumstancial cues

It is assumed that the driver perceives what others drivers are going to do by combining two sets of cues which together determine the construction of intentions: *categorisation cues* and *circumstancial cues*.

The *categorisation cues* are independent of the interaction between the car driver and other people. These are examples of natural categories (Dubois et al., 1993), constructed by subjects in the course of their driving experience and existing prior to the driving situation itself :

- *Formal rules*, such as the highway code ;
- *Informal rules* "a car does not remain behind a truck for long;
- *Stereotypes* that relates both to types of vehicle (slow, fast, taxi, truck, ambulance, scooter, and so on) and the types of driver (passive, aggressive, old, young...)

The *circumstancial cues* are dependent on the interaction and are constructed during the activity. They relate to the behavioural parameters of other driver, and the subject only takes them into account during the interaction process. A first set of variables, characterising the situation at a moment T (indicators, type of traffic, position of a vehicle at that moment, serves to construct a second set of inferred variables (trajectory, speed, a driver's previous behaviour) :

- *Variables at the moment T* : indicators (direction indicators, flashed headlights, brake lights, and so on) ; destination (when one finds oneself in an orientation zone, one infers the destination of someone else on account of his position) ; position on the road (in the middle or to one side) ; traffic characteristics (dense, fluid, stop-start) ; road infrastructure (insertion zone, orientation zone, exit zone) and alterations along the way (road-works, accidents) ;

- *Inferred variables* : trajectory ; speed (in relation to others or to oneself, in terms of slowing down or speeding up) ; a driver's previous behaviour ; lastly, a driver's behaviour must sometimes be conjectured with reference to the behaviour of a third party.

A single cue does not enable one to predict the action someone else will take. Only a combination of several elements can make that possible. The characteristics of the situation thus determine the significance of each of the elements the subject is liable to take into account. It is the probability of a change of state in each of those elements that leads the subject to pay more attention to some than to others. Since the situation is not stable, uncertainty persists. In order to justify his anticipation of another's behaviour, the subject views each cue he uses in terms of a state, or a development, or both at the same time. Those cues do not solely relate to what is happening at the time. To predict the action taken by someone else, the subject may refer back to the characteristics of a previous interaction situation that could have occurred a long time ago.

Not all the cues used to predict the action another user intends to take relate exclusively to the vehicle he is driving. They may characterise one or several other vehicles. Users thus anticipate another driver's behaviour by understanding the actions of a third vehicle.

Inferring others' intentions: what mechanisms come into play?

In this part of the study, our aim is to see if the 10 subjects impute the same or different intentions to other users when interpreting their behaviour in similar types of driving situation. These intentions are expressed verbally as predictions of a driver's behaviour, such as: "I think the blue car is going to stay behind the truck." The first case is *agreement*, when the 10 subjects make the same predictions of the driver's intention. We thus compare the combination of cues used to infer the other's intention. What is interesting is that an intention can be construed through various combinations of cues. The second case is *disagreement*: subjects make different predictions about what is going to happen. We have tried to describe the different forms of *disagreement*.

Agreement among subjects when inferring others' intention

In half the cases, the subjects infer that another person will take the same action on the basis of a combination of cues, whether common or not. We thus distinguish between central cues and secondary cues.

Apart from a very few exceptions, central cues, which are shared by several subjects if not all of them, always fall into the category of circumstantial variables. Hence, interaction-dependent variables dominate in guiding people's inferences of other driver's future behaviour. The most frequent of those central cues are, in descending order, the distance between two vehicles, their speed, and then their course and their position. Secondary cues, which relate to categories and rules, but also to traffic density and the infrastructure, are more dependent on the subjects and their own past history. They pinpoint the predicted action by confirming the orientations afforded by inference. Where there is agreement, the same inference may be the result of combinations of a certain number of central cues or of completely different cues.

Disagreement among subjects when inferring others' intention

External divergence : Faced with an identical interaction situation, subjects may make different inferences about the actions that others will take. In some cases, these inferences share common cues. Their combination, or their association with other cues, thus leads to conflicting predictions of others' behaviour. But differing inferences may also be due to a different perception of the same situation. Speed or distance may be gauged differently. Hence, while giving priority to the same cue (s), subjects opt for different actions. In this case, the conflict is a result not of a difference in the importance attached to the cue but to a subjective difference in the assessment of the situation.

Internal divergence : A same subject may experience uncertainty in predicting another's future action. The combination of conflicting cues prevents him from inferring a single future action. In some cases, in order to remove the uncertainty, he draws on a categorial variable (or stereotype). The latter result in expectations that may or not be corroborated by the circumstantial variables. Depending on the consistency of this combination, the driver predicts with a greater or lesser degree of certainty the other's future action.

One might think that these two forms of divergence could be sources of malfunction. In the absence of explicit information or a clear understanding, drivers may be led to make an erroneous prediction of another's future action, or to fail to adopt the most appropriate behaviour during the interaction because of the lack of readability of the other's action.

CONCLUSION

In conclusion, we will briefly discuss the respective contributions made by the two studies to the analysis of the mechanisms at work in the management of interactions between drivers. The first study, based on an in-depth analysis of drivers' activity in real driving conditions, enables us to identify a number of critical interaction situations, to define the situational demands that flow from them and characterise the mechanisms by which drivers manage their interactions with other users. The second study was devoted to deepening the analysis of these mechanisms and constructing models of them. It enabled us to identify the cues that are the starting-point for the prediction of other people's actions. We found that there were substantial differences from one subject to another when inferring others' future behaviour. These differences are likely to result in critical situations (mistaken expectations and/or belated detection of critical interactions, reducing the available margin for resolving them). These results, revealing drivers' difficulties in managing their interactions with others, should contribute to identifying ways and means for facilitating those interactions.

REFERENCES

Carré, J.R., Lassarre, S., & Liger, M. (1984). Fluidité-sécurité : Faut-il choisir ? *TEC*, n° 66, 14-19.

Dubois, D., Fleury, D., & Mazet, C. (1993). Représentations catégorielles : perception et/ou action ? *In* A. Weill-Fassina, P. Rabardel, & D. Dubois (Eds) *Représentations pour l'action*, Toulouse : Octarès.

Faverge, J. M. (1970). L'homme agent de fiabilité et d'infiabilité. Ergonomics, 13, 3, 301-327.

Leutzbach, W. (1974). Collisions par l'arrière. *12ème semaine internationale d'étude technique de la circulation et de la sécurité*. Thème VI : Collision en chaîne. Belgrade 2-7 Septembre.

Mazet, C. (1991). Perception et action dans la catégorisation. Le cas de l'environnement urbain et routier. Thèse

Malaterre, G. (1990). Error analysis and in-depth accident studies. *Ergonomics*, *33*, 9/10, 1403-1421.

Monseur, M., & Malaterre, G. (1969). Prise de décision des conducteurs aux carrefours. *Le Travail Humain*, *32*, 3-4, 217-238.

Mundutéguy, C. (1999). L'autre dans la conduite automobile. Prédiction de ses actions à venir. *Assises Européennes de Psychologie Appliquée aux Transports*. Angers, 16-19 juin 1999.

Risser, R. (1985). Behaviour in traffic conflict situations. *Accident analysis & Prevention*, *17*, 2, 179-197.

Saad, F., Delhomme, P. & Van Eslande, P. (1990). Driver's speed regulation when negotiating intersections. *In* M. Koshi (Ed.), *Transportation and Traffic Theory*, Elsevier.

Saad, F. (1996). Driver strategies in car-following situations. In Gale (Ed) *Vision In Vehicles V*. North Holland, Elsevier.

Saad, F. (1997). Managing car-following situations : an analysis of drivers' strategies in real driving conditions. In *Proceedings of the 13th triennial Congress of the International Ergonomics Association*, 433-435.

Van Eslande, P., Alberton, L., Nachtergaele, C., & Blancher, G. (1997). Scénarios-types de production de l'erreur humaine dans l'accident de la route. *Rapport de recherche INRETS* n°218.

RÉSUMÉ

GESTION DES INTERACTIONS ENTRE CONDUCTEURS D'AUTOMOBILES : UNE DIMENSION ESSENTIELLE DE LA SÉCURITÉ DE LA CONDUITE

Alors que de nombreux travaux ont été consacrés à l'analyse des interactions entre conducteurs et infrastructure routière, les modalités de gestion des interactions entre conducteurs ont, à notre connaissance, été assez peu étudiées. La gestion de ces interactions joue pourtant un rôle important pour la sécurité de la conduite dans la plupart de situations routières. Les recherches présentées ont pour objet d'approfondir l'analyse des modalités de gestion des interactions entre conducteurs et d'en proposer une première formalisation. Elles doivent, à terme, contribuer à l'identification de modes de communication susceptibles de faciliter la gestion de ces interactions.

Mots clés : Conduite automobile, conduite en file critique, gestion des interactions, prédiction des actions, incertitude.

Groupware Task Analysis in practice: a scientific approach meets security problems

Reinard van Loo, Gerrit C. van der Veer, Martijn van Welie

Dept Computer Science, Vrije Universiteit, De Boelelaan 1081 A,
1081 HV Amsterdam, The Netherlands. gerrit@cs.vu.nl

ABSTRACT

Groupware Task Analysis (GTA) is applied for the redesign of a commercial security system. The problem is the confidentiality of the knowledge on the task domain. The system is the actual security systems in use in companies, AND, obviously, none of them is eager to have details of its security management situation and security procedures being made available to outsiders, even if these are employed by a company that designed their system. The paper describes the problems in applying GTA and shows ways to cope with the restrictions involved in analysing safety critical systems in industrial design projects.

Keywords Design, groupware, GTA, security systems, task analysis

GTA - AN APPROACH TOWARDS DESIGN OF COMPLEX SYSTEMS

GTA (Van der Veer, Lenting, and Bergevoet 1996) is an approach to the early phases of the design of complex systems (task analysis, global design) that integrates ethnographic techniques and interaction analysis, analytic techniques from HCI, formal task modelling, and design space analysis. Analysing a complex system means analysing the world in which the system functions, or the "context of use". GTA deals with the context of use of a system in the broadest sense. In traditional literature on task analysis from the HCI (Human-Computer Interaction) mainstream, the focus is mostly on users, tasks, and software. CSCW (Computer Supported Collaborative Work) design approaches often focus on analysing the world first of all from the point of view of the (physical and social) environment. GTA, integrates viewpoints from the different approaches.

GTA Viewpoints on the Task World

1. **Agents** are considered in relation to the task world, hence, we need to make a distinction between actors and the roles they play. Moreover, we need the concept of organisation of agents.

- The label *actors* mostly refers to individual people, either individual or in groups. In situations where modern information technology is applied, actors will sometimes be non-human agents, or systems that comprise collaboration between human agents and machine agents.
- *Roles* indicate classes of actors to whom certain subsets of tasks are allocated, by free choice or as the result of the organisation. By definition roles are generic for the task world. More than one actor may perform the same role, and a single actor may have several roles at the same time.
- *Organisation* refers to the relation between actors and roles in respect to task allocation. The organisation describes the agent structure in the task domain. Delegation and mandating responsibilities from one role to another is part of the organisation, as is the way roles are allocated to actors.

2. We focus on the structural as well as dynamic aspect of **work**, hence, we take 'task' as the basic concept, and 'goal' as an attribute. We make a distinction between tasks (which have a goal) and actions in the 'classical' HCI terminology, and, moreover, we will elaborate task structure and the structure-related concepts of protocol and strategy.

- *Task* can be identified at various levels of complexity. Complex tasks may be split up between actors or roles. Users' unit tasks (and systems' basic tasks) may be decomposed further into (user) *actions* and (system) events, but these can only be understood in a frame of reference created by the corresponding task, i.e., actions derive their meaning from the task.
- The *task structure* will often at least partially be hierarchical. Task structures are not always known by single actors, mainly when different roles are involved in performing different subtasks. On the other hand, performance on certain subtasks may influence the procedures for other subtasks, which makes task flow and data flow structures relevant.
- *Protocols* indicates actual 'rules' as turn out to be applied for decomposing tasks, to be distinguished from 'rules' that may be stated explicitly in instructions which are sometimes not actually followed. Protocols may be situated, i.e., the environment and the presence of actors with certain roles may constitute conditions

for protocols to be triggered. *Strategies* indicate structures that can be considered protocols used mainly by experts or typically preferred by them. Strategies will be role related.

3. Analysing a task world from the viewpoint of the **situation** means detecting and describing the environment (physical, conceptual, and social) and the objects in the environment. Object description includes an analysis of the object structure.

- Each thing that is relevant to the work in a certain situation is an *object* in the sense of task analysis. Objects may be physical things, or conceptual (non-material) things like messages, gestures, passwords, stories, or signatures.
- The task *environment* is the current situation for the performance of a certain task. It includes actors with roles, conditions for task performance and for strategies and protocols, relevant objects, and artefacts like information technology that are available for subtask delegation. The *history* and temporal structure of relevant events in the task situation is part of the actual environment. The environment features as condition for task structures (inclusive protocols and strategies as far as these are situated).

The first phase in task analysis applying GTA: analysing the current task situation (Task model 1)

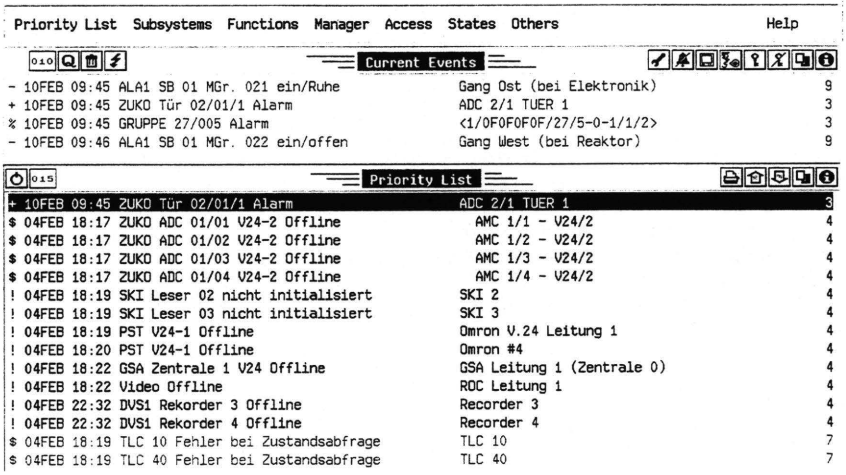
Task analysis may include different activities: (a) analysing a "current" task situation and modelling this (Task model 1), (b) envisioning a task situation for which information technology is to be designed (Task model 2), and (c) specifying the semantics of the information technology to be designed. Our current experience concerns phase (a) only, so we will not go into details of the other phases of task analysis activities. In many cases the design of a new system is triggered by an existing task situation. Either the current way of performing tasks is not considered optimal, or the availability of new technology is expected to allow improvement over current methods. A systematic analysis and modelling of the current situation may help formulate design requirements, and at the same time may later on allow evaluation of the design. Collecting task knowledge for task model 1 will need to be done by applying a variety of techniques. These techniques may be derived from "classical" HCI design methods like TKS (Johnson, Johnson, Waddington, and Shouls 1988) and MAD (Scapin and Pierret-Golbreich 1989). Often we will also need hermeneutic interpretation techniques (Van der Veer 1990). Additionally we will need to adopt ethnographic techniques that have been used for groupware design and CSCW, like interaction analysis (Jordan 1996). Actual techniques to collect the raw data for GTA normally include ethnographic observation, camera recording of interactions, and tape recordings of interviews

EUTERPE, a Task Analysis Tool

Our task analysis environment EUTERPE(van Welie, van der Veer, and Eliëns 1998) has been developed to aid the process of task analysis in GTA. EUTERPE was developed because task analysis is still an activity that needs support. After the task world has been modelled it is up to the analysts to interpret the task model and find out where causes of problems can be found or where there is room for optimisation of the work. A task model that can describe the task world including co-operative aspects and that allows some form of analysis could improve the task analysis process and outcome. Preferably the analysis of the task model should be done (semi-) automatically, thereby reducing the required effort of the analysts. However, performing a formal analysis of a task model requires a formal representation of the task model that is suitable for doing an analysis, especially for analysing co-operation. The task model therefore needs to be based on a task analysis theory that recognises the co-operative aspects of the task world. Although a formal analysis can be the basis for analysis it is not on the level analysts prefer to work. Hence representation tools can effectively hide the formalism and provide means to assist in analysing the environment that is being studied. In addition, a tool can also provide more structured ways of doing task analysis. EUTERPE, based on GTA, supports formal analysis both on a logical and a visual level. It uses formal representations internally while offering graphical representations to the user.

AN INDUSTRIAL DESIGN CASE: THE SECURITY SYSTEM

Currently the GTA-Method is applied at the Austrian Research Centre Seibersdorf (ARCS) mainly for the redesign of a security system marketed by Philips Industry, Austria. The main problem in this case is related to the confidentiality of the knowledge on the task domain. It is the actual security systems in use in these companies that is the basis for our knowledge of the task domain. Obviously, none of these companies is eager to have details of its security management situation and security procedures being made available to outsiders, even if they are employed by a company that designed their system. This paper describes the problems encountered in applying GTA in a real life redesign situation and shows ways to cope with the restrictions involved in analysing safety critical systems. Additionally, this paper derives additional suggestions on how to adjust task analysis for this type of industrial design projects. Securing large objects like factories, museums, banks or airfields is no small feat. Monitoring and controlling areas in these objects is done with the help of movement-detection systems, video camera systems, access control systems, fire detector systems, elevator control systems etc. In practice, all the information from these (sub-)systems is led to a control room where human operators have to respond appropriately to (alarm)signals. Keeping an overview on the building status becomes almost unmanageable in complex combinations of sub-systems. To support the operator in monitoring the state of the secured object and to integrate the different subsystems into one system, the sCalable Security System (CSS) has been developed. The CSS integrates the information flow from -and to- all subsystems in one central computer system with a generic (graphical) user interface (see figure 1) available on several workstations.



```
- 10FEB 09:45 ALA1 SB 01 MGr. 021 ein/Ruhe        Gang Ost (bei Elektronik)        9
+ 10FEB 09:45 ZUKO Tür 02/01/1 Alarm              ADC 2/1 TUER 1                   3
% 10FEB 09:45 GRUPPE 27/005 Alarm                 <1/0F0F0F0F/27/5-0-1/1/2>        3
- 10FEB 09:46 ALA1 SB 01 MGr. 022 ein/offen       Gang West (bei Reaktor)          9
```

```
+ 10FEB 09:45 ZUKO Tür 02/01/1 Alarm              ADC 2/1 TUER 1                   3
$ 04FEB 18:17 ZUKO ADC 01/01 V24-2 Offline          AMC 1/1 - V24/2                4
$ 04FEB 18:17 ZUKO ADC 01/02 V24-2 Offline          AMC 1/2 - V24/2                4
$ 04FEB 18:17 ZUKO ADC 01/03 V24-2 Offline          AMC 1/3 - V24/2                4
$ 04FEB 18:17 ZUKO ADC 01/04 V24-2 Offline          AMC 1/4 - V24/2                4
! 04FEB 18:19 SKI Leser 02 nicht initialisiert    SKI 2                            4
! 04FEB 18:19 SKI Leser 03 nicht initialisiert    SKI 3                            4
! 04FEB 18:19 PST V24-1 Offline                   Omron V.24 Leitung 1             4
! 04FEB 18:20 PST V24-1 Offline                   Omron #4                         4
! 04FEB 18:22 GSA Zentrale 1 V24 Offline          GSA Leitung 1 (Zentrale 0)       4
! 04FEB 18:22 Video Offline                       ROC Leitung 1                    4
! 04FEB 22:32 DVS1 Rekorder 3 Offline             Recorder 3                       4
! 04FEB 22:32 DVS1 Rekorder 4 Offline             Recorder 4                       4
$ 04FEB 18:19 TLC 10 Fehler bei Zustandsabfrage   TLC 10                           7
$ 04FEB 18:19 TLC 40 Fehler bei Zustandsabfrage   TLC 40                           7
```

Figure 1: The CSS user interface (menu bar, current events list and priority list)

The general concept of the CSS Operator interface is, that it reports every sub-system message that is defined by the manager as ‘relevant to monitor’. Examples of such messages are ‘Door contact XXX at ZZZ Alarm’, ‘Fire detector YYY at ZZZ Alarm’, ‘Door contact YYY at BBB Malfunction’, ‘Intrusion subsystem YYY offline’ etc. When a message enters the CSS a short beep is generated. The interface shows two lists: one first in-first out *Current Events* list where all relevant messages are listed and a *Priority* list. The messages in the *Current Events* list are supposed to be acknowledged by the operator one-by-one in the order of them entering the system. The list also contains ‘information’-type messages (like ‘Fire subsystem XXX online’) and messages that are important to remember. After acknowledging these last kind of messages they are automatically transferred to the *Priority* list, where they only disappear after the situation causing the message is somehow resolved. Little knowledge was available about how the current system is used and what kind of problems the users have with the system and/or User Interface in performing their work. To gain more insight on this topic, the first phase of the Groupware Task Analysis (GTA) method was performed. The analysis focussed on the use of the system by the end-users (i.e. operators, porters and system-managers) in their actual use-environments (factory, chemical factory and office buildings were visited). Goal of the analysis was to gain insight in the current use of the system and to propose directions of change/extensions to improve the (practical) usability of the system. The GTA framework, method and tools were highly usable for this.

TASK MODEL 1 - ACCESSING SOURCES OF TASK KNOWLEDGE ON THE CURRENT SYSTEM

Participating observations and semi-directed interviews were first carried out at two sites that seemed relevant. Getting the necessary co-operation of the managers and employees of the visited sites took some effort. Explaining the goal of the visits by telling some characteristic cases and ensuring that the resulting information remained confidential proved to be helpful in this process. At the end of the visits most managers and employees were enthusiastic about the fact that finally someone took the time to listen to them, and took their grievances serious. Due to the sensitive nature of the work observed (Security control rooms inside the objects to be secured) camera recordings (as usual in observations) and tape recordings of interviews were (almost) not possible. This did not impede the data gathering process.

Crucial to the acceptance of GTA in the project-organisation were the following factors:

- **External funding**. The analysis of the CSS was carried out in the context of the European OLOS Project (EC Human Capital Programme CHRX-CT94-0577). This made it possible for ARCS to get to know GTA with low financial risks. The results were the main ground for extending the appointment of the researcher at ARCS, outside European funding, to apply the method also on other projects.
- **Complementary expertise**. The results of the analysis made visible a gap in the acquired expertise in the project organisation: The direct translation of context-of-use characteristics in design information.
- **GTA being a method**. It offers a more solid base to work and communicate from. One does not start from scratch. Also, in getting support for gathering your data it is also very handy to mention that you are using a method. In a way, one abuses connotations of the method-concept like 'objectivity' and 'being systematic'; notions that make a technology oriented organisation more receptive

In successfully accessing the sources of task knowledge the following factors were crucial:

- **Confidentiality**. When performing several levels of confidentiality have to be kept. As with ethnographic observations in general, by being an observant on site one enters a social setting with its own rules, hierarchy etc. For the observed and interviewed operators it is important that what is observed does not impede the relation with management. Depending on the situation one stresses this point more (where a tension between management and operators is present) or less (where management and operators have a more open working environment). For the managers it is more important that the results of the analysis are treated with care and remain confidential. During the visits all these concerns could be resolved.
- **Not directly evaluating people**. On a more personal level, assuring that it is not a evaluation of the performance of the operators or managers in performing their job makes the people present in the use-organisation more co-operative.
- **Making clear Focus analysis**. The goal (improvement of the CSS) and focus (People performing work in an environment) of visits at use-organisations have to be made clear to all participants. This focus in connection with improving system design has to be explained further because normally these two are not considered in connection with each other. A characteristic case where it is clear that the system was not attuned to the use-organisation on the work-level can help (see for instance, (Neumann 1994)). It opens up a way of thinking about work and the role of the system in it where the participants can easily relate to.

PROBLEMS OF COLLECTING AND MODELING TASK KNOWLEDGE OF SECURITY SYSTEMS

First of all we had to choose which sites should be visited for our data collection. The following criteria were used in selecting sites:

- **Type of secured object**. The CSS security system is in use at over 70 sites including banks and factories. Every installation is tailored to those aspects that are most important to maintain or observe: At banks, for instance, maintaining control over who entered the building, and the monitoring of some secure areas are the primary concerns. At Factories, however, the primary goal is mostly the prevention of fire. These differences are also reflected in the work performed in control rooms.
- **Number of subsystems used**. The number and kind of subsystems differ considerably between sites: From simple (2 subsystems) to very complex (10-15 subsystems, 60.000 detectors etc.). This complexity also has its consequences for the work being performed.
- **How long in use**. The time the CSS has been installed also differs per site. At some sites, the system is rather new to the operators while at other sites the operators have accustomed themselves to the system. This also has consequences.
- **In flux or not**. Some sites have a very stable configuration, at other sites constant changes are being made to the configuration: new subsystems are added, construction work changes the layout of the building requiring the replacement of cameras, detectors and the like.
- Due to the limited time, we choose to start with as much **complexity** as possible. Two different kind of secured objects (office building and factory) were visited. Each site had many subsystems, the CSS had been installed for some time, and the configuration was in flux -so that the operators were able to compare the old with the new situation. Later on two other sites with low complexity and low flux were paid a short visit to verify and to get some feedback on observations done at the first two sites.

Using the chosen task knowledge gathering techniques (Ethnographic observation and semi-structure Interviews) in a straightforward way was severely impeded by situational and practical circumstances. Therefore, a more dynamic/improvisational mode of use of these techniques was taken. Videotaping during the observations was not allowed for security reasons. Therefore a ‘paper-and-pencil’ method was adopted for ‘recording’ the observed work. During the observations this was not perceived as an impediment to the analysis. It does however have consequences for how one works. Usually, video recordings are, among other things, used to analyse specific interesting actions: Interaction analysis. Not being able to do this afterwards calls for interaction analysis *on the spot*. A positive side effect of this is, that immediate feedback can be obtained from the observed persons. The cycle then becomes one of *observe – analyse – feedback*: One observes an interesting event (while not known yet, or along the way identified as a *hot spot*, see (Jordan 1996)), writes down a short task structure, and then the observed person can be asked about the event. This gives the possibility to check the interpretation just made. This also calls for another mode of observing, more active on the observer part: more intrusive. It is noted that the observed people do not mind this, as long as the observer stays close to the work. Tape-recording the interviews gave fewer problems. At most of the sites it was allowed to use the tape recorder to record the interview. At one site it was, however, not allowed. Not being able to record an interview was perceived as a severe impediment. On top of this, the interviews with operators at that specific site were only allowed in the presence of the security-manager, making the interviewed person less open about their work. Also, the manager only had the time to be present at one interview, so only one interview in total could be held. However, remaining questions could be resolved during a later short visit to the office building.

The interviews were done using a semi-structured approach, the so-called 'What-How-Why' role-specific interview as proposed in (Sebillotte 1995). This is used to gain insight into the task structure as explicitly seen by an actor. Using this method the problem of 'talkative' interviewees emerged: The semi-structured interview method had different results with different people. The interviewed operators were very short and not precise in their answers, leaving little room for the interviewer to prod deeper. During the observations, however, the same operators had much less difficulties in talking about their work. Whenever something happened and one was asked about it detailed, general, descriptions could be given. The interviewed managers, however, could describe their work in more detail. It is felt that the type of person and the situation in which the person was interviewed had a great influence on the quality and depth of the interview. During task modelling (using EUTERPE) it was found that the **interviews** resulted in more *complete* information about the work being carried out. The overall structure and personal interpretation of the work became clearer. The **observations**, however, resulted in more *detailed* and *correct* information about the work being carried out. In the task analysis these prove to be highly complementary to each order: correcting, filling up the blank spaces and giving structure to observed actions.

ADAPTING THE GTA APPROACH AND TECHNIQUES

Depending on the situation, different data gathering techniques are needed. What GTA lacks at this moment is a 'toolbox' of data gathering techniques with which one can start an analysis. These tools should act as building blocks or inspiration for new data gathering processes, where the situational constraints determine the choice of technique or combination of techniques. Such a toolbox should describe several techniques and, more importantly, their characteristics should be made explicit. What is it that the technique actually does? What kind of data does it produce? In the PD and HCI fields many different techniques can be found which are regarded as useful for performing task analyses. The criteria of application of such techniques in our approach should be that the results of the technique add to a better understanding of the work in terms of the GTA framework. The analysis as performed is the first step in the GTA method and was directly aimed at a kind of 'evaluation' of the studied system. This influences task modelling and the communication of results. One influence was to have the surface structure of models made to incorporate aspects that had to be made explicit for communication or for understanding by the analyst himself. An example was the use of new constructors: The usual task constructors were extended with 'conditional' versions of it: CALT, CPAR etc. This was meant to make clear that the execution of the tasks could be interrupted by a higher priority task, after which the old task was picked up again. This was such an important aspect of the observed work that the surface of the task model should signal this. During analysis the GTA *framework* proved to be the greatest help. It worked as a kind of 'check-list' to focus attention to things that matter in performing tasks. During analysis it was felt that *methodical* support was lacking, though there seemed not to be a real need for this. More needed is an openness on the side of the analyst in learning to understand why people do something. One has to be able to '*just enter a room*' and start observing something, which might turn out useless or useful for the analysis. One has to be able to deal with the open-ended-ness associated with a more 'ethnographic' style of performing analysis. That the GTA method leaves enough room for this without loosing the focus for system design as witnessed in the framework and later stages is regarded as a strong point, not as a weak.

CONCLUSIONS - GTA IN INDUSTRIAL PRACTICE

The promise of GTA as a system design method could not be validated in the case of our field investigations. In this project our help was called for late in the system design process and changing the process in terms of the

GTA design approach was not possible. This is, however, not considered as a problem: as part of a larger design team and project culture the results of the analysis stated in the GTA framework (work, people, Environment) and the translation of these findings into system design are incorporated in the project. Especially the 'translation' is crucial, and this turned out to be part of the core benefit of the GTA method in our case. The outcomes of the analysis were translated in concrete recommendations for interface and system design. This way the results were regarded very beneficial to the overall project. One acts as an intermediary between users and system designers and one has to be able to speak both languages. If one stays too far away from either the conceptual world of the users or of the designers the method won't work. Reactions to the GTA method are very positive and money is spent to carry on the work within the GTA framework. However, what is still lacking is more 'awareness' in the company of what it means to perform task analysis's at the customer and how it should be integrated in the already present system design culture. To the majority of traditional IT-personal and organisations it is radically new. It takes time to let this awareness grow by (partly) applying the method in current projects. Time is needed to move the application of the method forward in new projects.

REFERENCES

Johnson, P., Johnson, H., Waddington, R. and Shouls, A. (1988), *Task-Related Knowledge Structures: Analysis, Modeling and Application*, in: Jones, D. M. and Winder, R., People and Computers IV pp. 35-62, University Press, Cambridge.

Jordan , B. (1996), *Ethnographic Workplace Studies and CSCW*, in: D. Shapiro, M. J. Tauber and R. Traunmueller, The Design of Computer Supported Cooperative Work and Groupware Systems , North-Holland, Amsterdam.

Neumann, P. G. (1994), *Computer-Related Risks*, Addison-Wesley, New York

Scapin, D. and Pierret-Golbreich, C. (1989), *Towards a method for Task Description: MAD*, in: Berlinguet, L. and Berthelette, D., Work With Display Units 89 , Elsevier, Amsterdam.

Sebillotte, S. (1995), Methodology Guide to Task Analysis With the Goal of Extracting Relevant Characteristics for Human-Computer Interfaces, *International Journal of Human-Computer Interaction*, vol 7, no. 4, pp.341-363.

Van der Veer, G.C. (1990) Human-Computer Interaction: Learning, Individual Differences, and Design Recommendations, Ph.D. Dissertation Vrije Universiteit, Amsterdam.

Van der Veer, G.C., Lenting, B.F. and Bergevoet, B.A.J. (1996), GTA: Groupware Task Analysis - Modeling Complexity, *Acta Psychologica*, no. 91, pp.297-322.

Van der Veer, G. C., van Vliet, J. C., and Lenting, B. F. (1995), Designing complex systems - a structured activity, *DIS'95, Symposium on Designing Interactive Systems*, ACM Press, New York.

van Welie, M., van der Veer, G. C., and Eliëns, A. (1998), Euterpe - Tool support for analyzing cooperative environments, *Ninth European Conference on Cognitive Ergonomics*, Limerick, Ireland.

RÉSUMÉ

L'ANALYSE DE LA TÂCHE COLLECTIVE ASSISTÉE EN PRATIQUE : UNE APPROCHE SCIENTIFIQUE RENCONTRE DES PROBLÈMES DE SÉCURITÉ

L'analyse de la tâche collective assistée (GTA) est appliquée à la reconception d'un système de sécurité commerciale. Le problème est la confidentialité des connaissances dans le domaine de la tâche. Le système analysé est celui qui est réellement en service dans des entreprises et, évidemment, aucune d'entre elles n'est désireuse de donner des détails sur sa gestion de la sécurité à des personnes extérieures, même si ces dernières sont employées par une société qui a conçu leur système. Cette communication décrit le problème d'application de GTA et propose des façons de gérer les restrictions que l'on rencontre dans l'analyse des systèmes critiques de sécurité dans les projets de conception industrielle.

MOTS CLÉS : Conception, Travail collectif assisté, GTA, Systèmes de sécurité, Analyse de la tâche

Session 5: Error analysis
L'analyse de l'erreur

CIRAS: Collecting human factors data in the UK rail industry

L. B. Wright, J. B. Davies, E. Courtney & H. Reid
Centre for Applied Social Psychology
University of Strathclyde
Glasgow
e-mail: L.B.Wright@strath.ac.uk

ABSTRACT

This paper describes the Confidential Incident Reporting and Analysis System (CIRAS) which has been running since September 1996 for a number of railway companies in the UK. The aim of the system is to collect reports from individuals (drivers, signallers and other safety critical employees) of near misses, incidents and error-enforcing conditions, which would not normally be reported through normal channels, and to use this information to enhance existing safety management systems. CIRAS is confidential and "blame free", and therefore staff can report not only technical failures, but also operator or human errors without fear of recrimination and discipline.

BACKGROUND TO CIRAS

CIRAS is the Confidential Incident Reporting and Analysis System currently being used in order to identify and deal with human factors problems on the railways in the UK. CIRAS was initially a response to the contribution of human factors (including human error and latent failures) to incidents, situations and near misses on the railways. An earlier background report by Vosper Thornycroft pointed out the role of human factors in the rail industry, and the importance of these has also been highlighted in other industries (e.g. the off-shore oil industry; nuclear industry). Furthermore, existing official reporting procedures are often associated with disciplinary action, and this distorts the nature of reports received. This is particularly true in the rail industry where, historically, relationships between workforce and management have sometimes been characterised by mutual mistrust and animosity, rather than co-operation. This results in a tendency for reports to become focused on technical failures and chance happenings (in psychological jargon, the reports tend to be strategic, defensive and "external") with the human element being virtually absent (Van Vuuren, W., 1998). In some instances, it may even be the case that a near miss or incident with no obvious consequences will be deliberately concealed (i.e. the person concerned feels lucky to "have got away with it this time") due to the perceived disciplinary implications, rather than being seen as something from which others could usefully learn.

The CIRAS system by ensuring confidentiality seeks to rectify this imbalance by producing human factors data that otherwise go undetected and unrecorded. CIRAS is also timely since it complements the privatisation of the rail industry in the UK, opening the door to new, more open management systems and changes in safety culture.

WHO RUNS CIRAS?

The system was pioneered and adopted in the first instance by ScotRail Railways Ltd (train passenger service) in collaboration with the Centre for Applied Social Psychology, University of Strathclyde, for a trial period of two years. Those two years have now expired and in the light of the success of the project, the system is now expanding. At the time of writing the system is being used by ScotRail Railways Ltd, Railtrack Scotland Zone, Great North Eastern Railway, Virgin Trains (North) and eight companies carrying out contractual work for Railtrack Scotland Zone.

CIRAS is a University based, non-commercial system, and therefore has no vested commercial interests. It is concerned with improving safety on the railways, with collecting a national data base of human error from which all can benefit, and in developing predictive models and methods of analysis.

Although the system is administered and overseen by Professor John Davies and his staff, a Steering Committee was established to direct the general running of the project, to monitor progress, and to provide guidance and advice as necessary. The committee meets every three months and is chaired by Helen Muir, Professor of

Psychology at Cranfield University and includes representatives from the participating companies, HM Railway Inspectorate, Railtrack Safety and Standards Directorate and the University of Strathclyde. Three Liaison Committees exist to direct the project at a local level and include representatives from the companies involved plus the Health and Safety Executive, and trades unions. These committees are responsible for compiling feedback to safety critical staff in the form of a three monthly journal, and for recommending actions on issues of safety.

HOW DOES CIRAS OPERATE?

In the simplest terms, safety critical staff voluntarily report safety concerns, unsafe actions and practices direct to CIRAS. Reports are followed up by a telephone or face to face interview where more information is sought including demographic data on time, place, date, length of shift etc. Interviews are tape-recorded with permission and fully transcribed. These reports are then disidentified, and formal feedback is provided on the issues and incidents raised at three-monthly intervals. The feedback takes two forms. Firstly, there is the CIRAS Journal which is published quarterly and circulated to all employees who are potential reportees to the system. This provides information on the types of things reported, gives a management response, invites comment and includes a new CIRAS form. It has been demonstrated that such feedback is necessary to ensure the success of any reporting system (see Lucas in van der Schaaf, Lucas & Hale 1991). Secondly, a more detailed report is sent to management. This gives a more precise human factors description, includes some data-analysis, and makes recommendations where appropriate. As with the CIRAS Journal, all information is completely disidentified.

Confidentiality

Although CIRAS guarantees confidentiality, this does not mean that reports can be anonymous. In order to verify reports and obtain more information, it is essential that staff provide a name and contact telephone number and/ or address. The confidentiality process begins at this stage. At no point will CIRAS provide management with details that would identify an individual.

With a system like CIRAS, confidentiality is clearly of the greatest importance. The system has to be totally trustworthy, and has to be seen to be so. A single breach of confidence or security would cause the system to lose all credibility (e.g. number of reports decreased dramatically in a chemical processing industry following management decision to discipline an individual - personal communication with van der Schaaf). Accordingly, the CIRAS database is kept in a burglar-proof and protected location; data are stored on a removable hard drive which is kept in a safe overnight; and the computer itself is 'stand-alone' and non-networked so it cannot be hacked. The original report form is returned to the reportee, no copies are made or retained and all identifiers are removed when data are entered onto the computer. Thereafter it is impossible for any record to be traced to any individual through the database. Finally, all data are coded in terms of event-number, and no names or locations are recorded.

NUMBER OF REPORTS TO DATE

CIRAS has been operating for 34 months and 356 reports have been received from safety critical staff. The current reporting population is 5,700 including the very recent addition of 3,000 maintenance contractors. For the two companies who have been involved in CIRAS for the longest period of time (total reporting population of 1,800), the reporting rate is 11.9%. The potential reporting population includes drivers, signallers, train and track maintenance staff and other staff working in a safety critical post (e.g. platform dispatchers).

HUMAN FACTORS

Human factors problems extend all the way from the operator to latent failures (see Reason 1990). Thus, within the CIRAS system, human factors are broken down into three major categories. These are a) proximal factors, or mistakes that occur 'at the sharp end'; for example, in the driver's cab, or in the signal box; b) intermediate factors, which include maintenance, supervision, rules and communications; and c) distal factors which include procedures, design, management decisions and general company ethos / workforce culture. Within each of these three categories, the CIRAS system breaks event reports down into a finer grained and more specific set of codes.

Proximal Factors

The CIRAS model includes skill based, rule based and knowledge based errors (Reason 1990 and Rasmussen 1986). These and additional sharp end factors are shown below, with examples of the type of error made by the operator.

Proximal code	Definition	Example
Attention	Lack of concentration leading to slips (physical) and lapses (mental/ cognitive)	Failed to stop at station as singing in cab
Perception	Inability to see or hear specific features	Unable to see signal due to foliage
Knowledge	Lack of knowledge/ inadequate or incorrect knowledge for task	Trainee unaware shunting procedure not authorised
Rule violation	Deliberate breach of rules or procedures	Not using electrical protection when necessary
Rule based error	Use of wrong rule/ procedure in a given situation	Using obsolete braking instructions
Skill based error	Execution error	Conductor opened door on wrong side of the platform
Fatigue	Tiredness/ fatigue influencing behaviour	Train delayed as driver got out of train to walk along platform

Intermediate Factors

Intermediate code	Definition	Example
Communication between staff	Failure of communication between front line staff	Driver failed to pass on relevant information to signaller
Communication from staff to management	Failure of frontline staff to communicate with managers/ supervisors	Driver failed to report incident to supervisor
Communication from management to staff	Failure of supervisors to communicate with front line staff	Manager did not provide feedback on reported incident
Rule violation	Deliberate breach of rules/ procedures by supervisory staff	Briefing not given prior to work commencing
Maintenance	Inadequate or absent repairs/ maintenance	Monitors not maintained
Training	Insufficient for task, not provided	Training not provided on rarely performed task (isolating doors)

Distal Factors

Distal code	Definition	Example
Top down communication	Failure of senior managers to communicate with staff	Failure to fully explain implications of restructuring
Procedures	Ambiguous, difficult to follow or absent	Procedure surrounding inoperative AWS open to interpretation.
Design of equipment	Equipment design not fit for purpose	Lifting equipment for removal of train doors inadequately designed
Culture	Attitudes of work force, macho-style, general company ethos	Performance before safety

It is important that any coding system should be understandable to the managers who receive information from it. Therefore, CIRAS also has a classification system for technical/ equipment failures which is specific to the railway industry, which includes in-cab/ on-train equipment and infrastructure hardware e.g. signalling systems.

The data are assigned to the above human factors categories by CIRAS personnel. An inter-rater reliability trial was carried out on 28 randomly selected reports. Agreement between two independent raters was 84.6%. Across the data base as a whole 1, 058 codes have been assigned. This approximates to three causal codes per report.

The proximal codes consist mainly of rule violations, perception and attention difficulties. The number of rule violations (10% of the database) suggests that staff are willing to report actions and events for which they are responsible. At the intermediate level, the largest categories comprise maintenance and communication failures. At the distal level, procedures and culture account for the largest number of codes assigned. The pie chart below shows the percentage breakdown of the accumulated coded data in all categories.

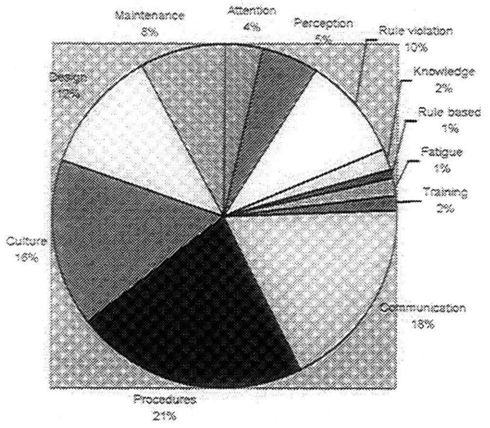


Figure 1: Human factors categories

POTENTIAL CONSEQUENCES

However, CIRAS also seeks to look at the possible *consequences* of near misses or failures, and attempts to discover why accidents and incidents are sometimes *averted* or avoided altogether. This is done through a set of 'consequence' and 'recovery' codes. At interview staff making reports are asked by the CIRAS researcher what could have happened had the situation not been recovered in some way. Responses are not prompted in any way by the researchers and are the opinions of those involved in the situation at the time.

Figure 2 below provides the potential consequences stated by employees at interview, for all reports received to date. The largest categories are potential fatality and potential SPAD (Signal Passed at Danger) which relate to the driving function (although a few of the potential fatalities relate to track and train maintenance). As the majority of reports concern Drivers it is not surprising that these categories are the largest. It is interesting to note that in 37 cases (just over 10%) of reports individuals felt that they would be open to disciplinary procedures for the issue or incident which they reported. While measuring the success of any reporting system is fraught with difficulty, the fact that staff are willing to report actions for which they could have been disciplined had they reported these events through normal channels leads to the conclusion that the system is worthwhile.

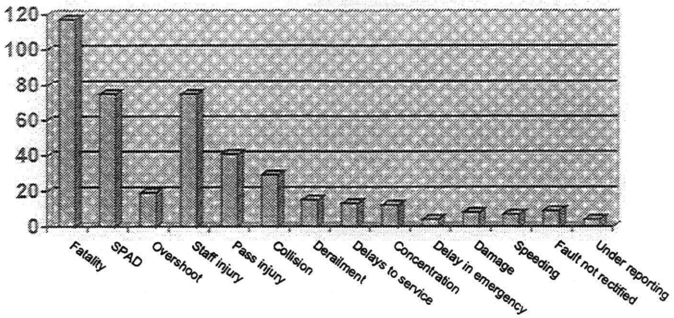


Figure 2: Potential consequences

RECOVERY STRATEGIES

An important aspect of any system such as CIRAS is to discover the reason why situations and near misses did not develop into fully fledged accidents. CIRAS seeks to do this via a set of broad recovery codes (see van der Schaaf 1988, and van der Schaaf et al 1991 for a discussion on the benefits of recovery). At interview, reportees are asked why they think the situation they were involved in did not lead to a more serious consequence. In 102 reports some form of recovery took place. The reasons given have been categorised into "procedures" i.e. situation retrieved by application of appropriate procedure; "individual action" i.e. situation retrieved by individual actions involving knowledge based or recognition primed decision; "technical" i.e. situation retrieved by automatic systems or technical barriers; "chance" i.e. situation retrieved by luck, guess work or trial and error. The classification of recoveries utilised by staff are shown in figure 3 below.

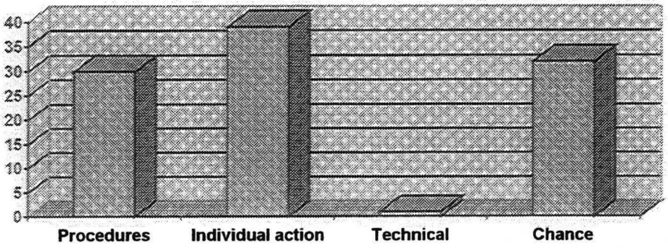


Figure 3: Recovery strategies

FUTURE DEVELOPMENTS

The companies involved are pleased with the results and types of incidents reported to date, and have indicated that the system has already saved money and highlighted situations which could have resulted in a number of serious incidents. However, further research is required on the reliability, validity and evaluation of the system and this is currently underway.

DISCUSSION

This paper describes the role of CIRAS as an important contributor to safety management systems on the railway network. Through reports from staff CIRAS aims to accumulate a database of human error from which all companies can learn. Within the railway industry there are both company and Railtrack led reporting systems – CIRAS does not seek to replace these systems, but to complement and add to the type of information that is currently gathered by these formal reporting systems. Since CIRAS was implemented in September

1996, the number of reports has steadily increased. The types of reports (near misses, personal involvement in unsafe practices) leads to the conclusion that the system is a successful and useful way of collecting information of a type that would not normally appear on official company reporting schemes. The benefits of receiving such information enable companies to target specific practices, improve ambiguous rules, and demonstrate their commitment to safety by making improvements where appropriate (e.g. enforcement of electrical protection despite increased work time required; rules regarding inoperative automatic warning system (AWS) clarified; shift diagrams changed following CIRAS reports of fatigue). This in turn can lead to an overall change in the company safety culture. Furthermore, issues raised through the CIRAS journal are used to highlight problems amongst staff at the official safety briefings. An on-going process exists to ensure that companies actively manage and respond to issues raised via CIRAS. CIRAS is in a position to integrate data from a wide number of companies and thus make a fundamental contribution to the reduction of human error on the UK rail networks.

REFERENCES

Lucas, D. A. (1991) Organisational aspects of near miss reporting. *In* Van der Schaaf, T. W., Lucas, D. A. and Hale, A. R. (eds) Near miss reporting as a safety tool. Oxford: Butterworth-Heinemann Ltd

Rasmussen, J. (1986) Information processing and human machine interaction. North-Holland, Amsterdam

Reason, J. (1990) Human Error. Cambridge: Cambridge University Press

Schaaf, T. W. van der (1988) Critical incidents and human recovery: some examples of research techniques. In: L.H.J. Goossens (ed.), Human Recovery: Proceedings of the COST A1 Working Group on Risk Analysis Human Error, Delft University of Technology, 13 October 1987.

Schaaf, T. W. van der Development of a near miss management system at a chemical processing plant. *In* Van der Schaaf, T. W., Lucas, D. A. and Hale, A. R. (eds) Near miss reporting as a safety tool. Oxford: Butterworth-Heinemann Ltd

Van Vuuren, W. (1998) Organisational failure: an exploratory study in the steel industry and the medical domain. (Ph.D. Thesis, Eindhoven University of Technology)

RÉSUMÉ

CIRAS : LE RECUEIL DE DONNÉES FACTEURS HUMAINS DANS L'INDUSTRIE BRITANNIQUE DU RAIL

Cette communication décrit le système d'analyse et de rapport confidentiel sur les incidents (CIRAS) qui est utilisé depuis Septembre 1996 par un certain nombre de compagnies de chemin de fer en Grande-Bretagne. Le but du système est de rassembler des comptes-rendus écrits par des individus (conducteurs, aiguilleurs et autres employés cruciaux pour la sécurité) sur des *near-misses,* des incidents ou des conditions favorables aux erreurs, qui ne seraient pas transmis par les canaux normaux. Cette collecte vise à améliorer les systèmes existants de gestion de la sécurité. Le CIRAS est confidentiel et garantit l'impunité. Par conséquent, les employés peuvent signaler, non seulement des défaillances techniques, mais aussi des erreurs humaines, sans crainte d'être réprimandés ou blâmés.

HUMAN FACTORS DATABASES FOR DESIGN AND SAFETY ASSESSMENT: THE CASE OF ADREP-2000 FOR THE AVIATION DOMAIN

P. C. Cacciabue, A. Perassi

European Commission,
Joint Research Centre,
Institute for Systems, Informatics and Safety
21020 Ispra (Va), Italy
pietro.cacciabue@jrc.it - arianna.perassi@jrc.it

ABSTRACT

In this paper, the importance and role of data sources and data collection in the domain of civil aviation is discussed from a theoretical and practical perspective. The data collection scheme, proposed in 1987 by the International Civil Aviation Organizations (ICAO), has been expanded for representing more accurately human contribution to accident causation, and a new classification scheme is proposed. This system is called ADREP-2000, and will be adopted in the future for classifying accidents. The theoretical paradigm proposed for ADREP-2000 is based on the well-known SHELL model, which presents advantages and drawbacks with respect to the simple structure of the previous classification scheme. These advantages and drawbacks have emerged from the analysis and codification of 9 major aviation accidents. A second important outcome from the study has been the consideration that ADREP-2000 is capable of representing an accident following the most modern theories of accident causation that tend to relate front line operators with the whole organisational structure.

Keywords: Human Factors data, accident databases, aviation accident classification, aviation accident analysis, root cause analysis, organisational studies.

INTRODUCTION

In modern technologically advanced domains and in aviation, in particular, the increase of the "human factor" element as cause of accident scenarios in the last two decades has been very remarkable and clearly emerges in many statistics (Rankin & Krichbaum, 1998). It is quite natural therefore that the design of human-machine interfaces and procedures, as well as system safety analysis and specially training, concentrate on human factors. The way to reach this goal is to devise methods for studying and improving Human-Machine Interaction (HMI) based on methodologies and theories that give consideration to relevant issues, such as cognitive and organisational aspects. A fundamental complementary element that contributes to the development of applicable and sound HF methods is the richness and quality of available data for performing analyses and calculation. Data and parameters are identified and generated in many different ways, which relate to analysis of tasks, to retrospective study of past events, to evaluation of work-settings and to collection of information by interviews and questionnaires.

In this realm of information search, structuring and retrieval, a crucial role is plaid by data contained in mandatory and voluntary reports on accidents, incidents, and near misses. In particular, the classification and databases derived from mandatory reports and the use of historical data and structured information are essential for the identification of root causes of accidents and for the development of proactive measures that aim at creating awareness and perception of risk and emergency management in case of accident.

To reach these objectives, a consolidated theoretical framework must support the ways in which data are collected and structured within a database. This implies that a number of human factors issues have to be taken into consideration in depth. These are mainly the cognitive aspects related to individual human performance, the socio-technical context in which the accident has occurred, the organisational climate, and the systemic and environmental conditions that may have favoured the occurrence of events. A widely accepted and applied theoretical framework that sustains these human factors considerations has been developed by Reason over a number of years (1990, 1997) and aims at framing the events that lead to an accident in a wide institutional context where all the components of an organisation are considered as possible contributors and causes to the accident. This approach has been preferred to other theories because of its simplicity and its comprehensive scope.

In the domain of aviation, the data collection scheme proposed by the International Civil Aviation Organizations (ICAO) in the late 80ies, i.e., ADREP-87 (ICAO, 1987), although inclusive of some human factors items, has not been considered sufficiently developed for representing appropriately the human contribution to accident causation in the light of the above considerations. Consequently, a new classification scheme has been proposed, i.e., ADREP-2000, which will be adopted in the future for classifying accidents. The system follows its predecessor in many respects, except for the human factors part, which has been totally revised, primarily for its theoretical background.

In this paper, the new ICAO classification scheme is shortly described focusing on its theoretical structure. Then the advantages and drawbacks that have emerged during a pilot study of classification of 9 real major accidents applying the new classification are discussed. The planned ameliorations and improvements to the classification that have been generated are also reported. Then, the ability of the new classification to give some insight on the organisational structure and liaisons that may have generated or sustained the chain of events is discussed. This will be done by showing how a specific accident can be codified by the ADREP-2000 scheme and the classification structure may, at the same time, reflect the accident sequence structuring of Reason's socio-technical and organisational theory.

THE CLASSIFICATION SCHEME ADREP-2000

ADREP-2000 Framework

Both ADREP reporting systems consider the accident as an *"occurrence"* and simply code all factors and findings, identified in the investigation along with the *"events"* to which they relate, usually structured in a temporal sequence. Most accidents consist of several inter-related *events*, such as "engine failure"-"undershoot". Each event is characterised by a *"phase"*, used to indicate at which stage of flight a certain event occurred. The *event-phase* pair must always be identified, as an event without its matching phase is of little value.

To describe events, a free number of *"descriptive factors"* can be entered for each event. To explain *descriptive factors*, any number of *"explanatory factors"* can be entered. Figure 1 illustrates how the two types of factors, i.e., descriptive factor and explanatory factor, relate to events.

More in detail, a *descriptive factor* describes *what* happened during an event by listing all existing phenomena. Descriptive factor consists of two parts: *"Subject"* and *"Modifier"*. The subject provides information on what was involved and the modifier gives the details. Human actions can in some cases represent a descriptive factor.

An *explanatory factor* explains *why* the event happened, when human performances or causes can be associated with the event. Explanatory factors only make sense when the involvement of persons or organisations has been ascertained. Explanatory factors consist of three parts: *"organisation"* or *"person"*, that indicates those involved or responsible; *"subject"*, that shows the area of involvement; and *"modifier"* that shows the nature of the involvement.

As mentioned above, the differences between ADREP-87 and 2000 concern mainly on the human factors aspects. Advantages of ADREP-2000 vs. ADREP-87 have already been discussed elsewhere (Perassi & Cacciabue, 1998) and here we will concentrate mainly on the issues related to the theoretical basis and the application of new scheme.

Amongst the various theoretical paradigms proposed over the years, the model originally proposed by Edwards (1972) and denominated SHEL, has been selected for supporting the development of a taxonomy and a related classification scheme for ADREP-2000. This model considers the human being (*Liveware*) as the focus of an interaction with other human beings (*Liveware-Liveware, L-L*), technical hardware system (*Liveware-Hardware, L-H*), technical software system (*Liveware-Software, L-S),* and socio-technical working environment (*Liveware-Environment, L-E).*

ADREP-2000 and Methodological Framework of Accident Causation

The process of classifying an accident according to ADREP-2000 can be defined as a "retrospective" analysis of an accident, as it is based on the considerations and findings of the investigations that follow accidents. The model that sustains the human factors part, as well as the overall structure of ADREP, are quite "old" as they are based on theories developed during the 70ies. Consequently, a critical process of review of the classification consists in the evaluation of its adequacy to represent an accident according to modern theories of accident causation. It this case, it has been considered whether an accident classified with ADREP-2000:

θ Reflects the combination of causal links depicted in Resason's theory of root cause analysis that looks into the whole organisational structure (Reason, 1990, 1997; Maurino, Reason, Johnston, & Lee, 1995);

θ Allows the evaluation and categorisation of the actual manifestations of errors made by humans while performing their tasks, in combination with the contextual and organisational factors affecting them; and, at the same time,

θ Respects the need to consider the temporal sequence of events, which is equally essential for a deep understanding of the accident evolution.

In short terms, Reason's approach for accident causation correlates human manifestations of errors (*active failures*) to: a) possible inadequate or erroneous practices, procedures, policies, and philosophies existing within a company; b) and/or to errors made in other circumstances, such as at design and maintenance level (*latent conditions/errors*); and c) to contextual conditions. This theory can be expanded so as to associate latent errors with system/hardware failures, in addition to human active errors. When focusing on the actual error manifestation these theories make reference to a detailed model of cognitive behaviour in order to catch the psycho-physical factors affecting overt behaviour.

Figure 2 shows the structure of the accident causation model based on the theory of latent/active failures of Reason, which has been overlapped with the structure of ADREP-2000/SHELL model, showing where and how the two frameworks meet.

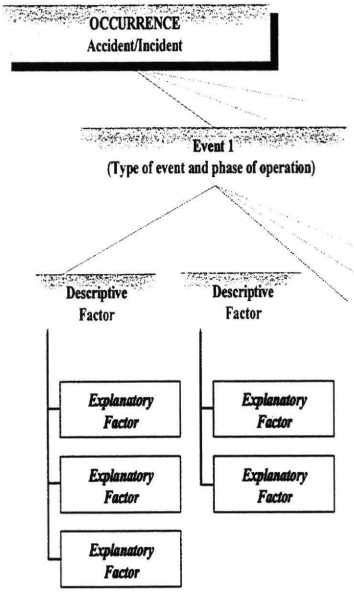


Figure 1. General structure of ADREP-2000 System

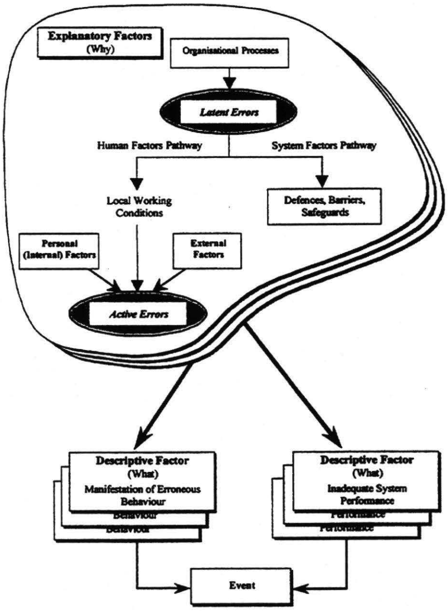


Figure 2. Accident Causation Model and Reporting Systems.

In particular, the following matching points can be observed:

1 The definition of *Event* is common to the two systems and represents the actual circumstance that has resulted from a combination of several human and systemic factors, situated in a certain time frame within the overall accident development. Consequently, an accident is made of a (temporal) sequence of *events* ("event time line").

2 The "Manifestations of erroneous behaviour" and the "Inadequate System Performances" identified in the accident causation framework correspond to the "Descriptive Factors" of ADREP-2000, as they both represent *what* has happened that led to the event.

3 The rest of the accident causation framework, i.e., the active/latent failure structure and the way in which latent failures can affect both human errors and system failures, as well as the influence of personal and external factors on human errors and system failures, is well captured by the SHELL model. In more detail:
   - "Personal internal factors" are characterised in ADREP-2000 by a complete set of items referring to Physical, Physiological, and Psychological conditions and to Workload management.
   - The remaining elements of the accident causation path affecting active failures, i.e., external factors, local working conditions, organisational processes can be found in many items L-E, L-H, L-L and L-S of ADREP-2000.
   - The "Defences, Barriers and Safeguards" line refers to the human contribution to system failures due to design errors as well as to inappropriate/poor/missing operational material and regulations (latent failure), which can be found in ADREP-2000 as L-H Interface.

In summary, both frameworks share the basic concept of accident causation and data reporting, that can be summarised as follows: The accident is the (temporal) sequence of *events* ("event time line"), each of which may be correlated to active human failures or system failures. These, in turn, can be linked to local conditions (usually random generated and mostly unpredictable) as well as to (systemically generated and thus predictable) organisational factors (latent failures/conditions) that may have been generated far back in time and within the organisation, during different stages of the system development. In this way it results possible to combine and correlate findings and outcomes of both accident analysis and reporting system in a structured format.

ASSESSMENT OF ADREP-2000

In order to evaluate the completeness of the information compiled into ADREP-2000 with respect to the findings of an accident investigation, we have considered 9 real accidents occurred over the past years in different locations, involving aeroplanes of different type and generation. Each accident has been studied by carefully reviewing the accident reports and findings of the investigations.

For each study, the most relevant *events* have been identified and further subdivided in specific findings, structured in "discursive" (not itemised) descriptive factors and explanatory factors, strictly related to the findings of the accident investigation. Each *event* and associated *descriptive* and *explanatory factors* have then been classified according to ADREP-2000. Details such as *type, phases, organisation/person, subject,* and *modifier* have been included as much as possible, utilising current supporting information.

The 9 events analysed are: Everglades accident, Zurich accident, Amsterdam Cargo-accident, Jeddah accident, Strasbourg accident, Cali accident, Los Angeles accident, Kegworth accident, and Langtang National Park accident. While the detailed accounting of the analysis for these 9 accidents is outside the scope of the present work and can be found elsewhere (Cacciabue & Perassi, 1999), the major results of this study with respect to the assessment of ADREP-2000 can be summarised as follows:

♣ ADREP-2000 is able to capture human factors aspects much better than the previous classification scheme. The main reason for this difference derives from the fact that ADREP -2000 is based on a reference model that enables to structure interaction between human and working context in a much clearer and complete perspective. This first conclusion could be drawn from the analysis of almost all nine accidents.

♣ Key elements that contributes to the clarity of *explanatory factors* in ADREP-2000 are the short comments associated with each item. A very important are support to the analyst is contained in these comments which make continuos references to other explanatory items that may have similar significance and, in some instances, may be more appropriate for the case at hand.

♣ The classification process with ADREP-2000 remains quite complex and cumbersome to complete, especially for the part relative to *explanatory factors*, even for a human factors specialist. As the analyst that is codifying an accident is usually an expert in aviation but seldom also a specialist in human factors, the risk of generating inappropriate coding remains quite high. For this reason there is a very significant need to develop a support tool that would provide appropriate guidance in the application of the taxonomy.

Discussion of a Case Study

In order to give an example of the work carried out, the Everglades case study will now be discussed in some detail, showing how the accident was analysed and classified and then how the accident causation model was superimposed to the classification.

ValuJet Flight 592 took off from Miami Runway and shortly thereafter the crew requested to return to Miami due to smoke in the cockpit. The flight 592 was vectored for a runway 12 approach. At 7207ft, descending at 260kts, the Flight Data Recorder stopped recording. Fifty seconds later ValuJet 592 struck a swamp with the nose pitched down and disintegrated. It was concluded that there had been a very intense fire in the middle of the forward cargo hold, which burned through the cabin floor at seat rows 5 and 6 on the left-hand side.
Investigations focus on a fire, possibly caused by oxygen generators carried in the cargo hold. The aircraft carried boxes containing 144 oxygen canisters and two MD-80 main wheel tyres in the forward hold.
The study and classification of the accident led to the identification of three major *events*: 1) Fire/ explosion/fumes, 2) Loss of control, and 3) Collision with terrain. For sake of brevity we will consider here on the analysis of *Event 1*. Figure 3 shows the classification of Event 1 by ADREP-2000 and contains also the numerical coding. If we focus on the analysis of contributing *Descriptive Factors 1* and *3*, which contain the most significant Human Factors issues, or *Explanatory Factors,* it is possible to reconstruct the accident causation root following the steps 1-3 described above (Figure 4 and 5).

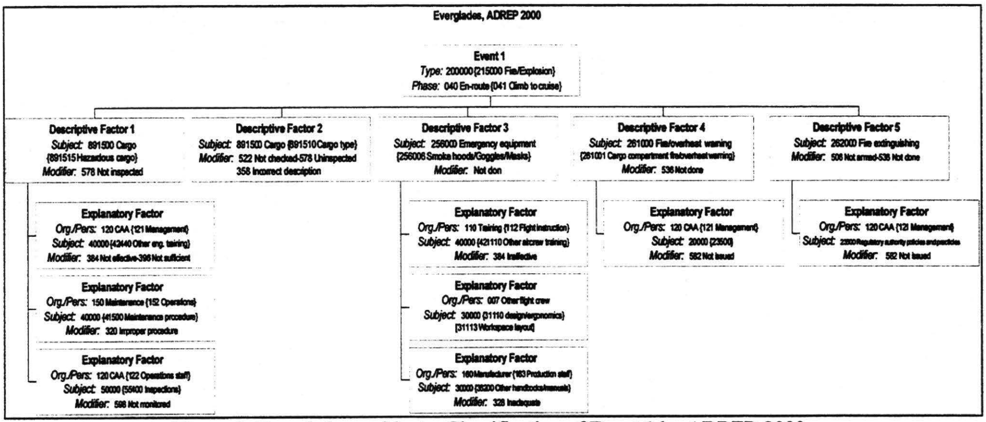


Figure 3. Everglades accident – Classification of Event 1 by ADREP-2000

In particular the following considerations can be made:

θ For Descriptive Factor 1, it becomes immediately clear how the lack of inspection of the hazardous cargo implies an *active failure* due to the maintenance personnel in properly performing procedures. This has been however, related to the conditions resulting from *latent errors* in the organisation, mainly at regulatory level, of insufficient training and lack of inspection requirement by the authorities (Figure 4).

θ For Descriptive Factor 3, the fact that the smoke masks were not don (*active error* and *descriptive factor*) has been linked to three latent conditions: a) lack of training; b) uncomfortable design of the workspace (the cabin space) and/or of the masks themselves (protective equipment); and c) inadequate manufacturer manuals production concerning the use of such equipment (Figure 5).

CONCLUSIONS

In this paper, the role of the data collection system called ADREP-2000 in the domain of civil aviation has been discussed, focusing on the ability of the system to capture the essential human factors contributors to an accident causation process. In particular, the adequacy of ADREP-2000 to represent an accident following the theory of Reason has been studied by analysing and codifying 9 major aviation accidents. A specific accident has been discussed in some detail to demonstrate the theoretical considerations.
The major conclusions that can be drawn from this exercise are that:

θ The ADREP-2000 classification system is able to capture the overall picture of links and correlation within an organisation; and

θ The data contained in a database that follows ADREP-2000 taxonomy allow a process of root cause assessment and evaluation throughout an organisation. This permits a reasonably accurate reconstruction of the sequence of events and human actions of an accident, without necessarily following a complicated and lengthy process of study of the complete accident report.

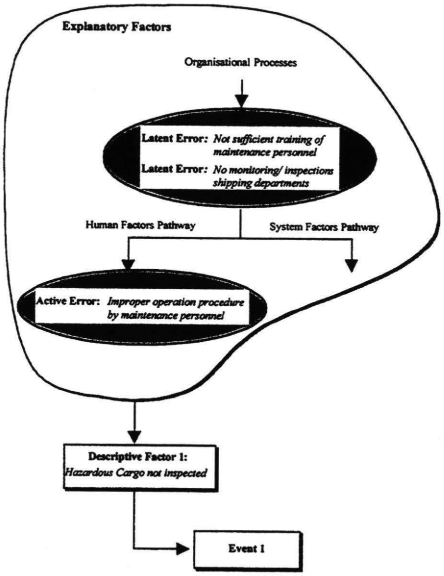


Figure 4. Event root analysis for Descriptive Factor 1 of Event 1 – Everglades accident

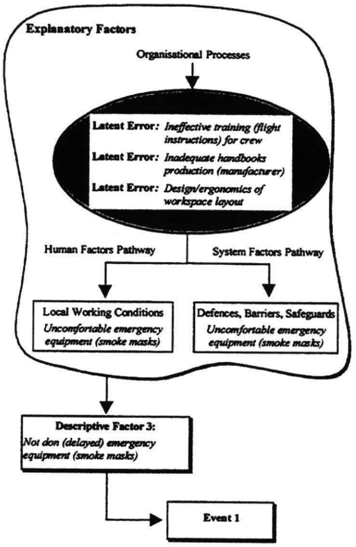


Figure 5. Event root analysis for Descriptive Factor 3 of Event 1 – Everglades accident

REFERENCES

Cacciabue, P. C., & Perassi, A. (1999). *Assessment of the ADREP-2000 classification system.* EUR Report to be published.

Edwards, E. (1972). *Man and machine: Systems for safety.* In Proceedings of *British Airline Pilots Association Technical Symposium.* British Airline Pilots Association, London. pp. 21-36

ICAO, 1987. *Accident/Incident Reporting Manual* Second edition - DOC 9156-AN/900. International Civil Aviation Organisation, Montreal, Canada.

Maurino, D. E., Reason, J., Johnston, N., & Lee, R. B. (1995). *Beyond Aviation Human Factors.* Avebury Aviation. Aldershot, UK.

Perassi, A., & Cacciabue, P. C. (1998). *Human Factors insight in databases of aviation accidents: comparing past and present classification schemes.* Proceedings of *17th European Annual Conference on Human Decision Making and Manual Control.* Valenciennes, France - December 14-16, 1998

Rankin, W., & Krichbaum, L. (1998). Human Factors in Aircraft Maintenance. "Integration of Recent HRA Developments with Applications to Maintenance in Aircraft and Nuclear Settings." June 8-10, 1998, Seattle, Washington, U.S.A.

Reason, J. (1990). *Human Error.* Cambridge University Press, Cambridge, UK.

Reason, J. (1997). *Managing the risks of organisational accidents.* Ashgate, Aldershot, UK.

RÉSUMÉ

LES BASES DE DONNÉES FACTEURS HUMAINS POUR LA CONCEPTION ET L'ÉVALUATION DE SÉCURITÉ : LE CAS DE L'ADREP-2000 POUR LE DOMAINE AÉRONAUTIQUE

Le schème de recueil de données proposé en 1987 par les organisations internationales de l'aviation civile a été étendu pour mieux cerner la contribution humaine aux causes des accidents. Cette extension, appelée ADREP-2000, sera adoptée à l'avenir. On présente ses intérêts et ses inconvénients à partir de l'analyse de 9 accidents aériens majeurs.

MOTS CLÉS : Données facteurs humains, Bases de données d'accidents, Classification des accidents aériens, Analyse des accidents aériens, Analyse des causes, Études organisationnelles

ANALYSIS OF THE HUMAN COMPONENT IN THE VEHICLE-ROAD SYSTEM

Gaetano Bosurgi (*) – Antonino D'Andrea (**) - Orazio Pellegrino (**)
bosurgi@ingegneria.unime.it staspc@tin.it

(*) Dipartimento di Costruzioni e Tecnologie Avanzate – Università di Messina
Salita Sperone, 31 – I-98166 Vill. S.Agata Messina ITALY - Fax +39 090 395022

(**) Istituto di Strade Ferrovie Aeroporti, Università di Catania, V.le Andrea Doria, 6
I-95125 Catania, ITALY

ABSTRACT

The trends on which modern road design are based have evidenced a lack of physical approaches in the definition of design standards. The need for analyses to contemplate a more complex system including the road user, the vehicle, the road and the environment is now well recognised. This article aims to evaluate analytically the risk of accidents due to aquaplaning through the exploitation of cognitive theories.

In particular, we wish to make a comparison between the results obtained using this methodology and traditional procedures such as those involved in risk analysis, like the Fault Tree. This should make it possible to establish effective utility and facilitate the future development of models accommodating more marked recognition of the human factor parameter.

We will also try to respond to the accusation that this type of analysis is highly subjective and, therefore, constitutes an unsatisfactory analytical approach to the problem by examining the instruments used in evaluating uncertainties which can also be used in conditions of limited available information.

KEYWORDS: Human Reliability – Cognitive – Driver Behaviour – Uncertainty Models

INTRODUCTION

Calculation of the chances of the feared event occurring is based on the definition of the random variable statistics involved which are, in the main: vehicle speed, driving experience and proficiency, tyre condition, road surface texture and thickness of water film, as Bosurgi, D'Andrea, & Pellegrino (1999) argue.

As is known, aquaplaning occurs when a significant thickness of water is present between the tyre and the road surface preventing adherence. This phenomenon is preceded by a spin down phase in which the anti-rolling forces exercised by the water film slow down the peripheral rotation of the wheel bringing it to a complete standstill. Spin down can vary by up to 100% at which point total aquaplaning occurs.

Di Mascio (1997) propose a Fault Tree of recognisable validity in which "Driver Behaviour" in emergency conditions, along with "Aquaplaning Conditions" and "Length of the tract affected by water film", decisively contributes to the triggering of the feared event, i.e. the risk of accident as a result of aquaplaning.

The study puts forward a sensitivity analysis which varies the probability of occurrence of individual factors while maintain all other factors at a value of 0.5. The results are shown in the graph in Fig. 1 where it can be seen that the components which most significantly influence the chances of an accident occurring are, in fact, the human behaviour and environmental components.

Fundamentally, from the literature generated in recent years, the need to include the human behaviour component in calculation procedures for road infrastructures has been fully recognised but at the same time methodologies have not been properly defined.

To tackle this problem adequately, it is necessary to have at one's disposal a comprehensive model which allows for examination of all the cognitive functions involved even at the cost of sacrificing a degree of analytical detail. The development of such a model is all the more necessary if we consider the myriad scientific studies of recent years and their definition of a number of different theories which have, however, centred only around specific aspects of driving activity rather than overall driver behaviour with the result that variations in peripheral conditions make the final result difficult to use in verifying such cases.

The procedure proposed in such articles involves illustration of the cognitive model and quantification of the probability that anomalous driver behaviour might occur in relation to the feared event. The modelling of uncertainties will be briefly described and guidelines will be given for future developments.

DEFINITION OF THE COGNITIVE MODEL

The cognitive model faithfully replicates the one proposed by Hollnagel (1998) in that it is recognised to include characteristics which make it appropriate for use in more in-depth analyses.

As is known this model aims to go beyond the usual sequential logic typical of traditional risk analyses by proposing a system of classification within which passage through the various groups occurs by means of antecedent-consequent links. The main advantage, among many others, is that there are no predefined pathways as, for example in a Fault Tree but rather a scheme which can adapt to the requirements of specific analyses and can be extended as function only of the interest and costs involved.

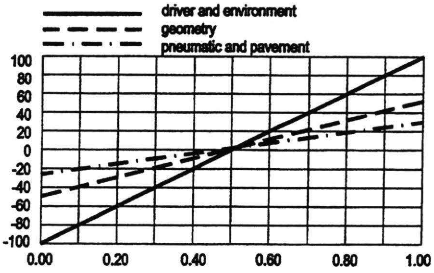


FIGURE 1 – Sensitivity Analysis

In defining the model the special peculiarities associated with in the various steps were evidenced of which the foremost is correct identification of context. Bearing this in mind a reference scenario was unambiguously defined within which road nature parameters were evaluated. The accuracy with which context and, therefore, peripheral conditions are defined is another characteristic which distinguishes such methodologies from traditional analyses. Indeed the latter have always considered the surrounding environment a possible transformer of the physical parameters involved in the phenomenon.

For example, the presence of a thickness of water film on the road platform is correlated to the problem of adherence between wheel and road surface, while no influence is assumed on driving behaviour.

For reasons of space the procedure proposed will be described briefly, so most of the intermediate results will be discussed without going into the greater detail that can easily be located in the bibliography should the reader require this further information.

USE OF THE COGNITIVE MODEL IN CONDITIONS OF AQUAPLANING

The procedure was applied to a real situation - a tract of State Road 113 located between 147 and 148 km. The choice of this tract was determined by its particular dangerousness which, over the years, has caused a series of accidents, some very serious.

It is constituted by a straight stretch of roughly 100 metres in length approximately half-way along which two secondary roads run into it. At the end of the straight stretch there is a curve to the left followed by a further straight stretch (Fig. 2). This curve is not characterised by a constant radius value but is made up of a succession of three circle arches each with a different radius. The road width varies from a minimum of about 6.40 m to a maximum of about 6.60 m in the region of the curve due to a slight widening. The transversal slope of the platform in the curved tract with the smallest radius (80 m) does not conform to the legal limits prescribed for efficient drainage of surface water and this is aggravated by the absence of lateral drainage channels. The path are all but inexistent as they are only about 0.20 m wide. The roadbed is delimited by the outer wall of an old building. Finally, the road surface is in a good state of repair both the surface and vertical traffic signs are insufficient.

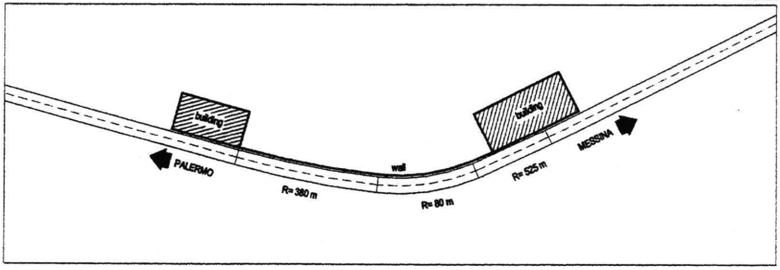


FIGURE 2 – Planimetry of Road Tract Under Study

Having outlined the situation, the first step is to define the surrounding environment in which action takes place by means of Common Performance Conditions (CPC) coefficients, which summarize the environmental peculiarities within the chosen scenario. We will thus have indicators for adequacy of organization, the working conditions the road user is subject to, the adequacy of MMI and operational support, the availability of procedures/plans, the number of simultaneous goals, available time, any circadian rhythm irregularities, adequacy of training and experience and crew collaboration quality.

For the sake of brevity we will avoid going into detail as regards the various CPCs and will only mention the values assigned to each of these as well as assessing indirect effects where coefficients are insignificant in such a way as to focus optimally on the most likely means of checking (Tab. 1).

These effects, according to conventions which can be varied as desired, make it possible to identify the most likely means of checking within the area in which action is taking place. The type of check derived in this example from knowledge of CPCs is of quite a high level and with it performance is strongly based on planning, on knowledge of the procedure to be followed and, in short, on strictly regularized behaviour.

TABLE 1 – CPC Scores

NAME CPC	LEVEL	EXPECTED EFFECT
Adequacy of Organization	Efficient	Not significant
Working Conditions	Compatible	Not significant (*Improved indirectly*)
Adequacy of MMI	Adequate	Not significant
Availability of Procedures	Inappropriate	Reduced
Number of simultaneous goals	Matching Current Capacity	Not significant (*Improved indirectly*)
Available time	Continuously inadequate	Reduced
Time of Day	Adjusted	Not significant
Adequacy of training and experience	Adequate, high experience	Improved
Drivers collaboration quality	Deficient	Reduced

The subsequent phases can be summed up by three fundamental stages leading to the final prediction as follows:

1. Development of the cognitive demand which is part of driving action.
2. Identification of probable errors within the cognitive functions.
3. Determination of the probability of overall action failure .

In order to define the kind of task the driver can expect, it is necessary to produce a criterion capable of defining and quantifying the cognitive activities involved.

The problem can be overcome, as in our case, using Hierarchical Task Analysis (HTA) but also by other similar methodologies which enable the main action to be subdivided into increasingly detailed steps distinguished from each other by one or more cognitive actions.

Indeed, the whole action was divided into three main phases described as *entrance, driving and exit* which constitute the first sequential type loop; entrance was, in turn, further divided into *check presence of other road users, check instrumentation, layout identification* and *speed check.*

At first sight, driving presents the same subdivision but, in fact, the cognitive activities involved will be different; in this iterative type loop the subdivisions are *check road user presence, check instrumentation, layout identification, speed check* and *direction check.*

The last loop, relating to coming out of the curve, is basically a return to initial conditions.

There is no need to linger over the finer points of this phase but it is certainly worth underlining that elementary actions are, nevertheless, highly dependent on easily measurable physical parameters. We can, therefore, predict whether or not these will be executed appropriately. These links can be simplified as follows:

Check presence of road users = f(traffic)
Check instrumentation = f(visibility)
Layout identification = f(adherence, evenness, centrifugal acceleration, visibility, plano-altimetric perils)
Speed check = f(initial speed, road user presence, plano-altimetric perils, vehicle condition)
Direction check = f(initial speed, road user presence, plano-altimetric perils, vehicle condition).

The availability of a data bank for these parameters, preferably expressed as probability density functions, would allow verification of criticalities. Nevertheless, here too there will still be computational difficulties due to the need for multi-dimensional integrations in calculating the distribution of marginal probability densities and relative covariances. Alternatively, approximate methods, already widely tried and tested, could be used but only when the functions which illustrate dependencies between these variables are known.

In the case under study, *visibility, plano-altimetric perils* and *speed* were beyond the permitted thresholds. This means that the functions involved were *layout identification, speed check* and *direction check;* that is to say that the existence of a good three variables which were over permitted limits affected a significant proportion of the cognitive functions the action entailed.

Cognitive activities, which at first sight might seem arbitrary, are in reality the result of exhaustive studies generally well accepted by the scientific community. For these reasons we will not include a complete list of these but concentrate only on those thought to be most significant for the purposes of the example under examination.

- *Diagnose:* the driver must interpret some signals, such as vehicle response and other parameters such as adherence conditions, presence of water film etc.
- *Evaluate:* it is necessary to take decisions using all the available information.
- *Identify:* Establish the identity of a system or sub system state.
- *Regulate:* Alter speed or direction of the vehicle in order to attain a goal.

These activities, which should help to identify the probabilities of failure, are closely linked the following cognitive functions: *observation, interpretation, planning and execution.*

The reason for differentiating between activities within a single cognitive function is that these, as already stated moreover, are connected to different phases within the main action.

The next step towards defining a cognitive demand profile involves linking cognitive activities to individual phases of the journey, as already illustrated using HTA. Table 2 also shows the links with the *observation (OBS), interpretation (INT), planning (PLA) and execution (EXE).*

TABLE 2 – Cognitive Demands Table for Driver Behaviour.

STEP	SUB TASK	ACTIVITY	OBS	INT	PLA	EXE.
Entrance	Identification of Layout	Diagnose	◆	v	v	◆
		Identify	◆	v	◆	◆
	Speed Check	Evaluate	◆	v	v	◆
		Regulate	v	◆	◆	v
Driving Progress	Direction Check	Regulate	v	◆	◆	v
	Identification of Layout	Diagnose	◆	v	v	◆
		Identify	◆	v	◆	◆
	Speed Check	Evaluate	◆	v	v	◆
		Regulate	v	◆	◆	v

It can be observed that a single phase presupposes the intervention of a number of cognitive activities but this is perfectly legitimate and, indeed, desirable if analysis is to be further refined.

At this point we are in possession of all the elements we need to sketch the histogram of the cognitive demand profile shown in Figure 3.

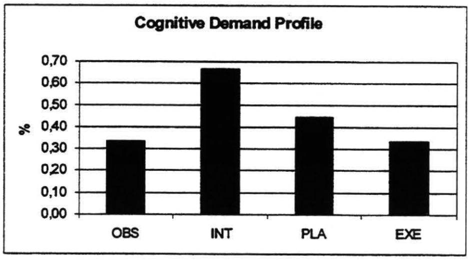


FIGURE 3 – Cognitive Demand Profile

The second phase in the prediction quantification procedure concerns the identification of probable errors in cognitive function processes. This is achieved by identifying all the possible antecedents, in limited numbers to avoid complicating analysis, while still contemplating all the probable sources of failure.

These modes of error will be outlined later but for the sake of brevity the deductive antecedent-consequent passage will be eliminated.

Naturally the distribution of cognitive function failures is dependent both on the cognitive demand profile and on the allotment of CPCs.

At this point it is possible to complete a table which relates cognitive activities to modes of failure where these have been further amplified within the four classes already defined. The results obtained will be ignored for obvious reasons of space. It is sufficient here to state that errors of *diagnosis* and *delayed interpretation* in the "Interpretation" category and errors of *execution* in the "Execution" category due to variables such as speed, direction and distance were the most decisive ones.

The third and last step concerns the determination of error probabilities. Indeed once probable failures of cognitive functions have been allotted for each sub task, it is possible to derive CFPs (Cognitive Failure Probability) for each of the error types so as to obtain a correct CFP value.

TABLE 3 – Adjusted CFPs for Cognitive Function Failures

SUB TASK	ERROR MODE	CFP NOMINAL	WEIGHT	CFP
F1.Driver Check	Faulty Diagnosis	0.2	5	2.5
	Decision Error	0.01	5	0.05
F2.Direction Check	Execution of Wrong Type	0.003	40	0.12
F3.Identification Layout	Delayed Interpretation	0.01	5	0.05
	Delayed Interpretation	0.01	5	0.05
F4.Speed Check	Decision Error	0.01	5	0.05
	Execution of Wrong Type	0.003	40	0.12

Nominal values can be derived from data gathered on an ad hoc basis or, as was done in the case under study, they can be found in literature.

The table below shows only the final phase of allotment of CFPs filtered by CPCs.

In our study, the lack of values dedicated to the road sector meant that the nominal CFPs were deduced from those available in literature.

To assess the probability of the final even occurring, the final CFPs will have to be incorporated in more complex calculation procedures like the Fault Tree or similar instruments. Future availability of these data will facilitate more specific application and, therefore, give more representative results.

MODELS OF UNCERTAINTY

The previous application needs, however, to be enhanced by more efficient analytical support in order to endow this type of procedure with a marked propensity for the resolution of road problems. The grey areas are located principally in the passages through the various HTA steps or between the various classification groups, so the final result is highly dependent on the capacities of the analyst. The difficulties of trying to apply mathematical models to the magnitudes involved lie in the fact that these are blatantly random and, furthermore, being intimately connected with human behaviour, we have no knowledge of the mechanisms that make these interdependent.

Having established the chance nature of these magnitudes, we need to propose back-up analysis procedures that make it possible to overcome the blatant approximations of deterministic mathematics.

Stochastic modelling is certainly the most widely used instrument and involves defining random variables and probability density functions. Within the stochastic context the main aim is to pinpoint the reliability of the system under examination, where by "system" we mean not only the mechanical components as a whole but also, when appropriate, the aspects relating to driver behaviour. Reliability, as defined by Rao (1992) is, therefore, the probability that the system will function adequately over a specific time range and under pre-established environmental conditions.

This definition contains four extremely important concepts which can be summarised as follows:

- reliability is expressed using a real number between 0 and 1;
- it must be established when a component should be considered a failure;
- operative conditions need to be established in that reliability can be greater or lesser according the the way the element under consideration is used;
- time of use needs to be specified.

Very often, however, experimental information is lacking and this means that the researcher relies rather heavily on the hypotheses formulated initially and , consequently, arrives at non-verified conclusions.

Thus, when information on these variables is fragmentary, it is necessary to use another approach in evaluating random factors. Elishakoff & Ben Haim (1990) proposed the convex modelling, which is simply a generalization of interval analysis. This theory does not define any measure of probability, but rather specifies sequences of events permitted as a function of the limited information available. In short, it is assumed that the random parameter x_j is constrained within a certain range $a_j'' \leq x_j \leq a_j'$ where this parameter can also be represented by a random function $x(t)'' \leq x(t) \leq x(t)'$ where the two extremes represent deterministic functions which delimit the variation range of $x(t)$.

These two examples, therefore, constitute convex models whose constraints contain an infinite set of values that the random variable may conform to.

With reference to the application previously proposed, three areas have been identified which are worthy of scrupulous further analytical investigation and which can be briefly described as follows:

1 Definition of context: the nine CPCs found in literature are simply environmental indicators which contribute to defining the peripheral conditions within which action is assumed to take place. For example the working condition factors, which is strictly dependent on the nature of physical conditions such as light, headlight dazzling, noise, activities which distract the driver, etc. could be monitored by defining stochastically one or more of these parameters in such a way as to make reliable positioning possible. The result which may even have originated from a simple semi-stochastic comparison with reference values should be located in one of the three or four classes into which the value of individual CPCs is located. It should be noted that this would simplify the assessment of indirect effects among CPCs.

2 HTA definition: As has already been seen, individual driving functions depend on a series of parameters such as traffic, evenness, centrifugal acceleration, plano-altimetric perils, visibility, initial speed, number of road users and vehicle conditions whose statistical moments can easily be described. Once the problem of defining probability density function has been solved for each of these parameters, the function which depends on more than one variable remains to be defined. This can probably be dealt with only through laboratory experiment using simulators. Once this is done, the expected results might be obtained, even if the density functions are not Gaussian, by means of approximate methods, as explained by Papoulis (1998). An initial development, rather than a simple choice made by the analyst, involves a sort of semi-stochastic check in which the stochastically characterised variable is compared with a deterministic value which can

be taken, for example, from reference norms.

3 Passage through classification groups: What was said for the two previous points remains valid, i.e. antecedent-consequent choice can be directed, if not defined absolutely, by knowledge of the random variables which distinguish human behaviour during driving activity. It is indispensable to identify the typical driver whose performance is being evaluated and whether a particular tract of road should be checked by testing it on the normal population, or whether, more particularly a specific class of driver is being monitored with reference, for example, to age, sight impairment etc. It is thought, however, that these objectives can be achieved using instruments such as convex analysis rather than by means of stochastic mathematics in that the information will, in any case, be unhomogeneous and, therefore, insufficient to describe the moments characterising the statistics.

From the above it can be inferred that there is every need to prefer calculation models that are not over-complex so as to limit analytical efforts; nevertheless, if the modelling of some random parameters represents the natural evolution of these procedures, one should tend towards representing functions which summarize most of the parameters involved and this can be refined and validated only by field work.

CONCLUSIONS AND FUTURE DEVELOPMENTS

Here we have tried to propose a more modern approach to the traditional calculation codes typical of road engineering in which, while acknowledging the fundamental importance of the human factor, have tended to reduce the phenomenon to a simple question of monitoring exclusively physical and mechanical parameters.

Thus, having evidenced the fact that human and environmental components represent the factors which most influence driving safety, we applied a general cognitive model recognised in literature to verify its adaptability to the requirements of the road sector and the chances of implementing within traditional calculation methods.

In the final phase we tried to overcome the danger of extreme subjectivity which characterises this procedure by identifying some areas in which the analysis of random factors can provide significant analytical back-up.

Future development in this sector is linked to the need to gather specific data on driver behaviour, isolating some of the more representative classes towards which further interest might be directed. For the same reasons the application of more complex cognitive theories requiring very high numbers of parameters still seems unlikely.

REFERENCES

Bosurgi, G., D'Andrea, A., & Pellegrino, O. (1999). *Aspetti relativi al comportamento di guida nel campo stradale* (Report n° 102). Catania: Quaderno ISFA Istituto di Strade Ferrovie Aeroporti – Università di Catania.

Di Mascio, P. (1997). *Analisi del rischio di incidente per Aquaplaning*. Paper presented at the Meeting on *La sicurezza Intrinseca delle Infrastrutture Stradali*. Roma, Italy, Feb.

Elishakoff, I., & Ben-Haim, Y. (1990). Probabilistic and Convex Models of Uncertainty in Structural Dynamics. *Proceedings of the 8th Intl. Modal Analysis Conference*, Orlando, 1487-1492.

Hollnagel, E. (1998). *Cognitive Reliability Engineering Analysis Method*. Oxford: Elsevier Science.

Papoulis, A. (1998). *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill.

Rao, S.S. (1992). *Reliability Based Design*. New York: Mc Graw-Hill.

RÉSUMÉ

ANALYSE DU FACTEUR HUMAIN DANS LE SYSTÈME VÉHICULE-ROUTE

Les courants sur lesquels la conception moderne des routes est fondée ont manifesté un manque d'approches physiques dans la définition des normes de conception. Désormais, la nécessité est bien reconnue d'analyses qui considèrent un système plus complexe incluant l'utilisateur de la route, la route et l'environnement. Cette communication vise à évaluer d'un point de vue analytique le risque d'accidents causés par l'aquaplanage, en exploitant des théories cognitives. En particulier nous souhaitons comparer les résultats obtenus en utilisant cette méthodologie et ceux auxquels conduisent les procédures classiques d'analyse des risques, telles que l'arbre des défaillances. Cela pourrait permettre d'établir une utilité effective et de faciliter le développement futur de modèles intégrant une reconnaissance plus marquée du facteur humain. Nous essaierons également de répondre à l'accusation selon laquelle ce type d'analyse est hautement subjectif et, par conséquent, constitue une approche analytique non satisfaisante du problème, en examinant les instruments utilisés dans l'évaluation des incertitudes qui peuvent aussi être utilisés dans des conditions de restriction de l'information disponible.

MOTS CLÉS : Fiabilité humaine, Cognition, Comportement du conducteur, Modèles d'incertitude.

Multivariate Analysis of Human Error Incidents Occurring at Nuclear Power Plants: Several Occurrence Patterns of Observed Human Errors

Yuko Hirotsu Kunihiko Suzuki Mitsuhiro Kojima Kenichi Takano

Central Research Institute of Electric Power Industry (CRIEPI),
2-11-1, Iwado Kita, Komae-shi, Tokyo, JAPAN
hirotsu@criepi.denken.or.jp kunihiko@criepi.denken.or.jp
kojima@criepi.denken.or.jp takano@criepi.denken.or.jp

ABSTRACT

This paper described the obtained results in respect of typical human errors together with causal factors and situational factors found commonly among analysed incidents at the domestic nuclear power plants (NPPs). CRIEPI has been conducting detailed and structured analysis of all incidents above reported during last 31 years and stored in a database. Firstly, all case studies were classified according to above factors. Next, trend analysis concerning frequencies of those was conducted. Finally, multivariate analysis was conducted. The result suggested that there were several occurrence patterns that were identified as how errors occur at NPPs.

Keywords

Nuclear power plants, human error, database, multivariate analysis, causal factor

INTRODUCTION

Total 885 incidents occurring at domestic nuclear power plants (NPPs) have been reported due to the requirements of the regulatory laws from the beginning of Japanese NPPs history. Among these incidents, 193 incidents have been identified as human error cases. While the number of human error cases (incidents / unit / year) is steadily decreasing year after year, the proportion of human error cases among total incidents is constantly about 20 percent.

Human errors could occur as a result of combinations of many causal and situational factors such as conditions of workplace (e.g. dark, narrow), human internal conditions (e.g. fatigue, inattentive), plant situations, and human interfaces. The observed interrelations of these factors, which suggest us typical patterns of how human errors have occurred at NPPs (occurrence patterns), will be useful information for preventing the recurrence of those having similar patterns by applying common countermeasures.

Cognitive approach analysis of observed human error cases have not been enough done because there were not enough detailed information on worker's cognitive process concerning each case. However, if above interrelations can be explained with a cognitive aspect, the results will be available for developing the research on psychological bias affecting human cognitive performance, which is in progress at CRIEPI.

The purpose of this study is to find above interrelations out of case studies. The following approach was employed in this analysis.

(1) Analysis of individual human error case using J-HPES (Takano, Sawayanagi, & Kabetani, 1994), which is a system to analyse and evaluate human error cases and to obtain data not only in symbols but also in textual descriptions such as causal factors, sequence of events, and proposed countermeasures;
(2) Classification of above data according to causal factors, situational factors and types of human errors into several definite categories;
(3) Trend analysis concerning frequencies of above categories;
(4) Multivariate analysis concerning above categories.

This paper summarised the obtained results in respect of typical human errors together with associated job categories, situational factors, and causal factors found commonly among analysed incident data during operational jobs and maintenance.

METHODS

Subject Data

Under co-operation with a governmental organisation, CRIEPI selected human error cases from all incidents reported during last thirty-one years, from the beginning of the first Tokai nuclear power operation till fiscal year of 1996, according to the detailed criteria (Nishijima, 1998). Total 193 incidents were chosen as human error cases occurring during jobs such as design, manufacture, installation, operation, maintenance, or management.

Analysis of Individual Case

J-HPES was applied as a method for analysing individual error case. This system comprises following three constituents.

(1) Implementation Procedure (a practical procedure on how to proceed analysis and evaluation): it comprises four stages, "Correct Understanding of Events" "Situational Analysis" "Causal Analysis" "Proposing Countermeasures". In Correct Understanding of Events, human errors are identified. In the Causal Analysis, the modified fault tree method is applied for initiating a search reaching down to the ultimate underlying causal factors.

(2) Evaluation Forms (formats in which to describe the obtained results): These forms comprise two parts; one is "Situational Analysis Forms" in which to record the results of associated situational data including consequence detection and task nature, second is "Summary Reports Forms" which provide summary with charts and tables.

(3) Evaluation Guide (a guide to aid personnel in efficiently performing their tasks in analysis and evaluation)

Analysis of individual case has been carried out by using computerised analysis system "JAESS" (Takano, et al, 1994) based on an event report, and analysed data have been stored in the J-HPES database. Each case was analysed under co-operation with plant experts, spending approximately eight hours.

Categories Used for Classification

Table 1 shows the categories used for classification. "Job Types" and "Situational Factors" had already been classified into categories in J-HPES data-base. Since it was difficult to classify the observed cases into performance levels (knowledge base, rule base, skill base), information process phase (direction(instruction), detection, judgement[†], implementation) was included in Situational Factors for analysis in a cognitive aspect. As to the other data, the categories were defined as a result of grouping operation that each factor having almost the same meaning was assembled using KJ method[‡]. Categories of "Causal Factors" were defined in two levels; general categories and detail categories.

Tab.1 Items and categories used for classification

Job Types (4 categories)
Human Error Types (9 categories)
Causal Factor (general:12 categories detail:42 categoris)
Situational Factors Information Proces Phase (4 categories) Plant Status(5 categories) Location (7 categories) Task Frequency (3 categories) Task Nature (3 categories) Task Repeatability (2 categories) Work Load (3 categories) Shift Schedule (2 categories)

Analysis Method

The correspondence analysis (Lebart, Morineau, & Warwick, 1984), one of the multivariate analyses, has been applied for a set of data classified above. When there are three and more variables (in this case, factors), the method facilitate one to grasp the interrelation between variables by replacing with less number of variables. This method can deal not only quantitative variables, but also qualitative variables, such as yes/no. It corresponds to a principal component analysis in case of analysing only quantitative variables.

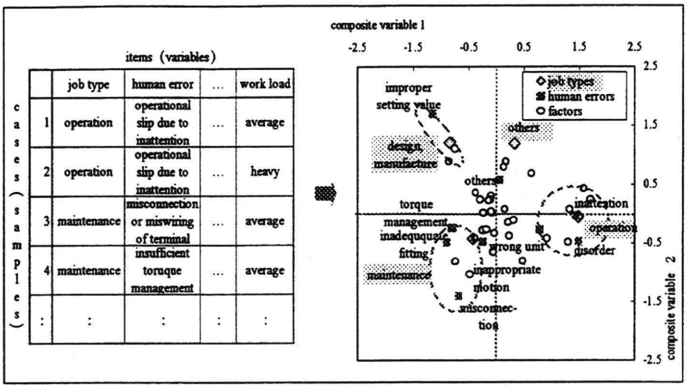
Figure 1 shows the concept of correspondence analysis. The left table in figure 1 represents a data set of human error cases classified into categories. When correspondence analysis is applied to the data set, scatter diagrams with two dimensions of new variables, which are the composites of original categorical variables, can be

---

[†] Judgement is used as a term comprising situation assessment, decision making, and response planning.

[‡] A method to put down fragmental information on cards, sort out them with the meaning, and categorize them.

obtained. In the scatter diagrams, original categories are plotted on a plane, and categories that have close interrelationship shall be gathered around gravity centres. The right diagram in figure 1 is an example of obtained scatter diagram. Categories enclosed with each dotted line especially have close interrelationship.



items (variables)

	job type	human error	...	work load
1	operation	operational slip due to inattention	...	average
2	operation	operational slip due to inattention	...	heavy
3	maintenance	misconnection or miswiring of terminal	...	average
4	maintenance	insufficient toruque management	...	average
:	:	:	:	:

Fig.1 The Image of Correspondence Analysis

RESULTS AND DISCUSSIONS

Categorical Classification

Human Error Categories

Table 2 shows the result of classifying all human error cases according to job types. This indicates that about 60% of those occurred during maintenance and about 20% during operational jobs. It is also indicated in the analysis of significat events at NPPs in USA by Institute of Nuclear Power Operations (INPO,1985) that the number of errors during maintenace was dominant.

Tab.2 Job types

Job types	Percent of total
Maintenance	58
Operational jobs	20
Design, manufacture	11
Others	11

As a result of grouping, observed human errors could be classified into major nine categories. Table 3 shows the eight categories, in which the most frequent error was *torque management when closing and tightening objects* which occupied about 15% of all errors classified. As to the other error categories, *wrong unit / train / component* and *operational deviation or disorder*, which might cause an severe impact on the NPP, occupied about 9% and 6%, respectably.

Tab.3 Human errors

Human errors	Percent of total
Torque management when closing and tightening objects	15
Misconnection or miswiring of terminals	9
Operational slip due to inattention	9
Wrong unit/train/component	9
Inappropriate motion of workers	8
Improper setting value	6
Operational deviation or disorder	6
Insufficient tightening or inadequate fitting object	5
Others	33

NPPs in Japan are required by laws and regulations to make an annual inspection of prescribed facilities and items. During these inspections, utilities carry out voluntary overhaul, inspection, and modification work for the purpose of preventive maintenance. Error categories such as *torque management when closing and tightening objects* and *misconnection or miswiring of terminals* occurred mainly in those work. In the case of aeroplane, assembling work occupied 70% of the cause of engine shutdown(Boeing,1994). These results indicate that there were no prominent and featured human errors even in the large-scale system such as NPPs or aeroplane, but tiny and common errors, which can happen at any industries and workplace, have resulted in incidents.

Causal Factor Categories (General)

Causal factors were also classified into twelve categories. Table 4 shows the categories in which the most frequent causal factor was *work practice* which occupied about 23% of all. Adding the percentage of *work schedule, work practice, work verification* and *change implementation* together, the categories concerning work cycle could reach up to 40%. As to the others, major categories were *human internal status* and *written communication*, that occupied about 19% and 13%, respectably.

Tab.4 Causal factors

Causal factors	Percent of total
Work practice	23
Human internal status	19
Written communication	13
Work verification	10
Interface design or insufficient demarcation	8
Supervision and management	7
Work schedule	5
Verbal communication	5
Work condition	4
Field condition	3
Training	2
Change Implementation	1

Specific Causes and Situational Factors Contributing to Each Human Error

Human Errors during Operational Jobs

Correspondence analysis clarified typical interrelationships between types of human errors during operational jobs and following items: Causal Factors, Situational Factors (Information Process Phase, Plant Status, Location, Task Frequency, Task Nature, Task Repeatability, Work Load, and Shift Schedule).

Figure 2 shows that different human errors are plotted separately, which means that they occurred from different causes and situations. The most typical human error occurrence patterns were identified as followings:

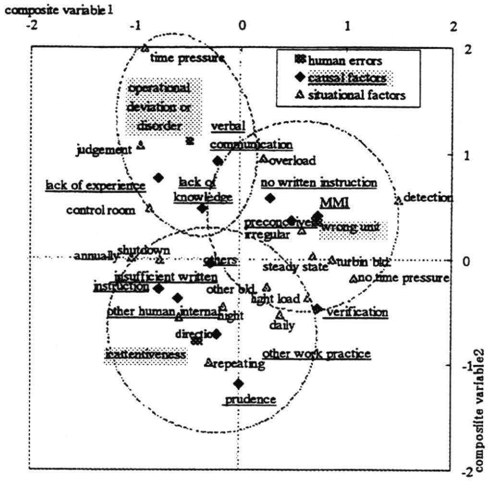


Fig2 Interrelation of human errors, causal factors, & situational factors during operational jobs

(a) Operational deviation or disorder

<Causal Factors>

[Verbal Communication], [Human Internal Status]-Lack of knowledge, Lack of experience

<Situational Factors>

Information Process Phase: judgement,
Occurrence Location: control room,
Task Nature: under time pressure

This interrelationship shows that *operational deviation or disorder* occurred in the phase of judgement and was caused by lack of communication among operators, lack of knowledge and lack of experience. It was likely to occur when an operator was doing the task with tight time schedule. Especially, it occurred in a control room

(b) Wrong unit / train / component

<Causal Factors>

[Interface Design or Insufficient Demarcation(MMI)], [Written Communication]-No written instruction or not using written instruction, [Human Internal Status]-preconceived idea, [Verification]

<Situational Factors>

Information Process Phase: detection, *Plant Status* steady state, *Occurrence Location*: turbine building,
Task Frequency: irregular, *Task Nature*: no time pressure, *Work load*: overload

This interrelationship shows that *wrong unit / train / component* occurred in the phase of detection and was caused by human machine interface such as inadequate layout and confusing display or tagging of valve numbers, lack of appropriate procedures or instruction, preconceived idea, and insufficient verification. It was likely to occur when an operator was not accustomed to a job. Especially, it occurred in a turbine building when the plant status was in full power, steady state.

(c) Operational slip due to inattentive

<Causal Factors>

[Work Practice]-Lack of prudence, Others (Unawareness of the situational change, etc)
[Written Communication]-Insufficient or ambiguous written instruction, [Human Internal Status]

<Situational Factors>

Information Process Phase: direction, *Plant Status*: under shutdown,
Occurrence Location: others (waste handling facility etc), *Task Frequency*: daily, annually,
Task Repeatability: repeating, *Work load*: light, *Shift Schedule*: evening / night

This interrelationship shows that *operational slip due to inattentiveness* occurred in the phase of direction and was mainly caused by lack of prudence, unawareness of the situational change, insufficient or ambiguous description in instructions or procedures, or confusion. It was likely to occur when an operator's arousal level was low as in night shift, or an operator was doing a task of less than average load, or a repetitive task. Especially, it occurred during an annual inspection.

Human Errors during Maintenance

Correspondence analysis told us typical interrelationships between types of human errors during maintenance and following items: Causal Factors, Situational Factors (Information Process Phase, Location, Task Frequency, Task Nature, Task Repeatability, and Work Load).

Figure 3 reveals that *inadequate fitting* and *insufficient torque management* occurr by similar causes. The most typical human error occurrence patterns were identified, such as:

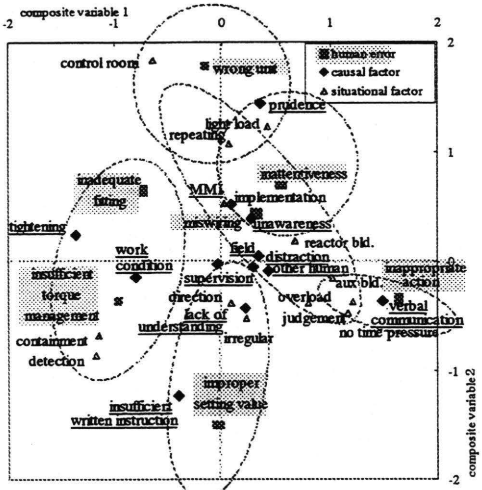


Fig.3 Interrelation of human errors, causal factors, & situational factors during maintenance

(a) Wrong unit / train / component

<Causal Factors>

Nothing special identified

<Situational Factors>

Occurrence Location: control room,

Task Repeatability: repeating, *Work load*: light

These interrelationship shows that *wrong unit/ train/ component* was likely to occur when a worker was doing a repetitive task or an easy task. Especially, it occurred in a control room.

(b) Improper setting value

<Causal Factors>

[Verbal Communication]- Insufficient or ambiguous written instruction,

[Supervision and Management],

[Human Internal Status]-Lack of understanding

<Situational Factors>

Information Process Phase: direction,

Task Frequency: irregular

This interrelationship shows that *improper setting value* occurred in the phase of direction and was caused by insufficient or ambiguous description in instructions or procedures, inadequate management, or worker's lack of understanding. It was likely to occur when a worker was not accustomed to a task.

(c) Misconnection or miswiring of terminals

<Causal Factors>

[Interface Design or Insufficient Demarcation(MMI)], [Field Condition], [Human Internal Status]-Distraction, Others

<Situational Factors>

Information Process Phase: judgement, *Occurrence Location*: reactor building, *Work load*: overload

This result shows that *misconnection or miswiring of terminals* occurred in the phase of judgement and was caused by inadequate design, unsuitable field condition such as unstable scaffold or insufficient lighting, or worker's distraction. It was likely to occur when a work load was heavy. Especially, it occurred in a reactor building.

(d) Insufficient tightening or inadequate fitting objects, Insufficient torque management

<Causal Factors>

[Work Condition], [Work Practice]-Inadequate tightening

<Situational Factors>

Information Process Phase: detection, *Occurrence Location*: containment

This interrelationship shows that *insufficient tightening or inadequate fitting objects* and *insufficient torque management* occurred in the phase of detection and were caused by work condition such as excessive amount of tasks or inadequate way of tightening. Especially, it occurred in a containment vessel.

(e) Slip due to inattentive

<Causal Factors>

[Work Practice]-Lack of prudence, Unawareness of the situational change

<Situational Factors>

Information Process Phase: implementation

This interrelationship shows that *slip due to inattentiveness* occurred in the phase of implementation and was caused by lack of prudence or unawareness of the situational change.

CONCLUSION

Concerning human error cases occurred from the beginning of the first domestic nuclear power operation, analysis of individual case, classification of data, trend analysis, and multivariate analysis have been carried out. It was suggested that there were several occurrence patterns identified. The followings are the major results of the analysis.

- **Observed human errors were tiny and common to any industries that occurred under job control and quality control.**
- **Even the type of human error is the same, the occurrence pattern of error is different according to the job type.** For example, *wrong unit / train / component* during operational jobs was caused by human machine interface, lack of appropriate procedures or instruction, preconceived idea, or insufficient verification. It was likely to occur when a operator was not accustomed to a job. However, errors comitted by maintenance personnel were likely to be arisen when a personnel was doing a repetitive task or an easy task.
- Human errors that would be important in a cognitive aspect are *wrong unit/train/component* and *operational deviation or disorder* during operational jobs, and *insufficient tightening or inadequate fitting objects, insufficient torque management,* and *misconnection or miswiring of terminals* during maintenance. Most of these were including detection and interpretation in information processes.

REFERENCES

Boeing (1994). *Maintenance Error Decision Aid, Boeing Commercial Airplane Group.* Seattle: Boeing Commercial Airplane Group

INPO (1985). An Analysis of Root Causes in 1983 and 1984 Significant Event Reports. *INPO 85-027.*

Lebart, L., Morineau, A., & Warwick, K.M. (Eds.). (1984). *Multivariate Descriptive Statistical Analysis : Correspondence Analysis and Related Techniques for Large Matrices.* New York: John Wiley & Sons, Inc.

Nishijima,Y. (1998). Human performance indicators. In J. Misumi, B. Wilpert, & R. Miller (Eds.), *Nuclear Safety: A Human Factors Perspective* (pp. 213-219). London: Taylor & Francis

Takano, K., Sawayanagi, K., & Kabetani, T. (1994). System for Analyzing and Evaluating Human-Related Nuclear Power Plant Incidents – Development of Remedy-Oriented Analysis and Evaluation Procedure–. *J. Nucl. Sci. Technol., 31,* 894-913.

Takano,K., et al. (1994). Analysis and evaluation system for human-related incidents at nuclear power plant. *J. Nucl. Sci. Technol., 36,* 1059-1067

RÉSUMÉ

ANALYSE MULTIVARIÉE DES INCIDENTS LIÉS À L'ERREUR HUMAINE DANS LES CENTRALES NUCLÉAIRES : PLUSIEURS CONFIGURATIONS D'OCCURRENCES D'OBSERVATIONS D'ERREURS HUMAINES

Ce texte décrit des résultats obtenus dans des centrales nucléaires domestiques sur des erreurs humaines typiques, ainsi que sur des facteurs causaux et situationnels couramment évoqués dans les analyses d'incidents. Le CRIEPI a conduit une analyse détaillée et structurée de tous les incidents rapportés pendant ces derniers 31 ans et stockés dans une base de donnée. Tout d'abord, toutes les études de cas ont été décrites selon les facteurs cités. Ensuite, une analyse de tendance a été réalisée sur les fréquences. Enfin, une analyse multivariée a été conduite. Les résultats montrent qu'il y a eu plusieurs configurations d'occurrences d'erreurs dans les centrales nucléaires.

MOTS CLÉS : Centrales nucléaires, Erreur humaine, Base de donnée, Analyse multivariée, Facteur causal

ROUTINE MENTAL PROCESSES IN NPP FIELD ACTIVITIES: TRACEABILITY- AND RELIABILITY-RELATED ASPECTS.

Alain NOIZET* & René AMALBERTI**

* Electricité de France (EDF), Pôle Industrie, Division Recherches et Développement, Groupe Facteurs Humains. 1 Avenue Général de Gaulle, 92140, Clamart, FRANCE. email : alain.noizet@edf.fr

** IMASSA, Département Sciences Cognitives, BP73, 91223, Brétigny-sur-Orge, email : rene-a@imaginet.fr

Abstract : The study stems from a request to study routine errors (confusions) as a cause of incidents in nuclear industry (Electricité de France, EDF). The paper deals with two points: first, the definition of routine, with a debate on the controversial approaches in the literature. Second, a critical approach of methodological aspects to study routines in the field. Various interview techniques, questionnaires, and a more experimental method based on simulation, are presented and criticized after tests in the field. Beyond these methods, the study ambitions to understand and model the role of routine in the dynamic control of cognition and risk management.

Keywords : Routines, Cognitive control, Cognitive compromise, Methodology, Simulation, NPP's field activities

The study aims at modeling the role of routines in NPP operators' cognitive activities. It stems from a request to study *confusions* as a cause of incidents in nuclear industry. In the past ten years, confusions of locations or devices have represented over 10% of the incidents stored in Electricité de France's (EDF) nuclear incident data base. Most of these incidents were considered minor, but could have induced more severe consequences in a different context. A preliminary study has (Noizet, 1998) shown that confusions mostly affected well-trained operators, and were made during routine activities. This is the reason why EDF's study progressively moved towards the concept of routines, and then of cognitive control of those activities.

The research is expected to go beyond the classic and simple relationship between routine and error, and to reach a better understanding of the role of routines in overall human cognitive performance. Studying routine activities in ergonomics faces three basic problems : (i) a problem of definition, (ii) a methodological problem to grasp routines in activity analysis; and, (iii) a problem of prevention. The paper addresses the two first points: routine definition and methodology.

DEFINITION OF ROUTINES

The concept of routine belongs to common language, and to various scientific contexts and disciplines (e.g., psychology, sociology, economics, and computer science). It is therefore not a surprise that the definition remains blurred and polysemic. After a brief reminding of the meanings of routine in common language and several contexts, this section gives emphasis to the psychological point of view. It ends with the author's definition of routines for the purpose of the EDF research.

Routines in common language: Common Sense

The notion of routine is generally understood as a brand of specific know-how observed in a frequently performed manual or intellectual activity. It allows rapid and loosely controlled performance. Common sense often lends a pejorative connotation to the notion of routine in considering that it implies 'a lack of intelligence' in the performance, often interpreted as an increased risk of unintentional errors. Moreover, it is often considered that the subject who acts routinely 'could have been more careful' and 'would have achieved better performance'.

Routines out of psychology: Routines in Computer Science

The concept of routine (or sub-routine) applies to sub-programs called by a main program. It does not bear any negative connotation. The use of routine in programming stress two advantages (i) better performance (better program architecture, faster run) and (ii) enhanced resource saving and economy (easier program readiness and debugging, re-use capacity).

Routines out of psychology: Routines in Economics

Routine, together with 'Convention' and 'Institution' (see Reynaud, 1998) is one of the three pillars of business organization. The notion of routine relates to two meanings: first, the routine is a behavioral model, regular, predictable, efficient and guided by norms; as such, the description of company's routines reflects the kernel of the company's know-how. Second, routine relates to a continuous dynamic learning model improving task performance by repetition and adaptation. To sum up, routines in economics are considered as conscious activities which represent the base of business knowledge and its evolution.

Routines out of psychology: Routines in Sociology

The use of routine in sociology is fairly rare. Nevertheless, when routines are quoted, the most frequent meaning is close to the one described in economics. Routine are seen as shared knowledge (among individuals) that organize group behaviors in normal life. Collective routines tend to reduce mutual checks, verbal comments, with a growth of implicit and tacit mutual knowledge.

Routines in psychology

The concept of routine is the last come of a series of neighboring concepts (levels of consciousness, automaticity, and skills) that have marked the psychological literature.

1/ Shiffrin &Schneider (1977) are among the first authors introducing clearly the distinction between automatic processes and controlled processes. Automatic processes are characterized by their speed (which explains their lack of access to introspection), and their lack of direct attentional control. Conversely, the controlled processes are slow and serial processes regulated by the subject. The global performance increases with automatic processes not only because of their higher speed of execution (as compared to controlled processes), but also because of the possibility of multiplying parallel automatic activities (automatic processes are considered resource free).

2/ Routine capacity is a natural outcome of leraning. Anderson (1982) suggests that expertise acquisition changes the status of declarative knowledge for that of operative procedural knowledge directly integrated in the performance procedures. Special attention is also paid to the change of cognitive control during the acquisition of expertise. Norman & Shallice (1986) insist on a permanent 'vertical' control of all parallel processes that play at the horizontal level of task execution. They specify that this vertical control decreases without ever disappearing altogether for routines. Reason (1984) and Rasmussen (1986) take up the same ideas: *Skill-Based Behavior* is never totally synonymous with no cognitive control.

3/ Logan (1988) proposes a global theory of automatisation alternative to the modal view (limited resources in attention). He explains the performance of automatic processes only by mean of a change of memory storage and memory access characteristics. Automatic processes are characterized by direct access to solutions stored in the memory (too rapid to give rise to conscious awareness), whereas controlled processes require recombining and elaboration work. Logan also lights on the distinction between *automaticity* and *skill* (*or routine*, terms with identical meaning for the author). Automaticity relates to basic processes and behaviors that can be proceeded with high speed, absence of effort, and cognitive autonomy; whereas skills and routines are presented as a coordinated combination of automated processes. The coordination/control among automatic processes is conscious, and depends on the declarative knowledge activated in context.

4/ Baars (1997) integrates most of these ideas in his theater metaphor. Certain actors are lit, while the rest of the stage and other actors "operate in the dark". This background consists of routines, which correspond to the unconscious components of voluntary action.

Human reliability and routine errors

Aside the literature on the mechanisms of routine, another important literature relates to the risk of routine error. Routine errors have given rise to an abundant literature on their activation conditions (Baars, 1992; Reason, 1984) and their prevalent detection mechanisms (Rizzo, Bagnara, & Viscolia, 1987; Sellen, 1994). Four points stand out in this literature:

1/ Routine errors represent 60 to 70% of all errors

2/ Routine is the producc of expertise. Accordingly, routine errors are more frequently made by experts

3/ Recovery of routine errors is better than that of any other errors (Rizzo & al, 1987). Norman & Shallice (1986) have suggested the existence of a parallel processor called SAS (Supervisory Attentional System) able to interrupt the performance of a routine in case of incoherence detection when intention and performance are compared. This supervisor is activated in the core of a routine when significant shifts in the procedural traces are detected, or takes advantage of conscious emergences, which coordinate activation of routine constituent automatisms and ensures the contextual adaptation of the activity (reaction to deleterious outcome).

4/ Routine errors concern all routinized actions, without exception, and without presuming of the seriousness of their consequences. Corrective modifications of the context may entail local effects (elimination of certain specific errors through interruption of the routine), but has no effect on the overall number of routine errors.

Moreover, solutions that could artificially counteract routine activation and maintain a high level of control (e.g. alerts, signals, mandatory checks, self-questioning attitudes) are themselves subject to learning and to secondary routinization.

Definition Adopted

The definition proposed for the research in progress rests on four principles derived from the review of literature:

- Mean-end principle: Routines are considered both as a cognitive mean (a cognitive tool to manage workload) and as an end (a structure of knowledge).
- Continuum principle: Any task can be carried out at various degree of routinisation. The degree of routinisation depends on the number and quality of cognitive controls invested in task completion. The pitch and level of control (simple detection of errors Vs conscious tactical elaboration of the next sequence) is part of a continuum dependent on the operator's choice (choice relying on the contextual model and cognitive arbitration of the moment). The organization of routine, and the cognitive points of control of automatic sequences, depend on the individual past experience of failure.
- Global stability principle: Any increase in the cognitive control of a task reduces the routine level of this task, and consequently entails an increase in the routinisation of competing tasks. When considering the global level of activity, the total amount of routines at anytime is important (maybe almost constant), but the distribution of routines considerably vary among the competing tasks composing the activity. Work activity is the result of an arbitration and a cognitive compromise between controlled activities and more routinised activities. This arbitration, reviewable at any moment of the performance, relies on a contextual model of the subjective perception of the control of the situation (Amalberti, 1996).
- Sufficient result: Routine is also the expression of a procedural knowledge, that can be triggered in context, with great efficiency in the domain of application. Routinised activities (or skills) are associated with a satisfactory and stable level of performance (reasonable performance stability in repetition, high speed , very few fatal errors, satisfactory result fitting the demand).

The authors intend to validate this model in the EDF' s study, with special emphasis on modeling the dynamic cognitive control of routine as part of the global control of activity. A specific interest is the description of the alerting scheme (contextual signals ?, cognitive signals?) that triggers the change of the dynamic control of routines. New guidelines for the design of the work environment, and for the training program, are expected as outcomes of the study. The research program is now entering its second year. The rest of the paper focuses on the methodological difficulties to grasp and model routines in complex activities.

METHODOLOGY FOR AN APPROACH TO ROUTINES

Intervention of a nuclear power plant operating crew member involves covering some distance on industrial sites, in several buildings and on several levels, in order to reach the site/premises where the equipment/objective to intervene on is located (reading of parameters, information taking, physical action on a system). In normal operation condition the operators are very familiar with these interventions (often carried out by a single operator) for having carried them out numerous times.

Few cognitive ergonomics methods allow to grasp and observe routines. Two reasons account for this methodological sub-specification:

- Poor literature on routine modeling in natural work activities. The almost sole focus is the negative outcome of routine behaviors at work and in everyday activities (routine errors).
- Limited access to sub-symbolic activities when using classic observation/ interview methods.

This part presents a critical overview of cognitive ergonomics methods (classical and more targeted) implemented in order to describe the routines of field operator.

Classical methods

- *Observation of natural situations*

Routines are often studied from their negative outcome (errors). However, routine errors inform more on the error mechanism and context (capture, lapsus, etc) than on the cognitive management of routine. One of the classical missing element in routine-error reporting is the description of what was the very nature of the global activity (context, evolution of control, competitive goals, stakes, tasks).

Out of the context of error, the natural grasp on activity prioritizes traces of the decision making processes and routines are chronically under traced (Kellog, 1982). For example, Dubuisson (1998) in his attempt to observe routines performed by regular travelers in a train station environment had no difficulty in identifying the behavior of the lost traveler (outside his routine) but encountered difficulty in distinguishing the regular travelers from the more casual travelers.

Besides, the observation of operators by means of video, or audio, tends to increases the level of control of their activity, and reduce the number of routine errors, as observed in the nuclear field (Noizet, 1998).

♦ *Think aloud protocols*

Routines only call for local involvement of the conscious system, at the most it involves being conscious of performing a task but without active focused attention on most sequences of activity (*conscience of determination* according to Norman & Shallice, 1986).

It is however possible with think aloud protocols to grasp the level of control of routinized activities. The study of the focus of consciousness (the object of verbalization) indirectly reveals what is not controlled during task completion, and also where are the emergences of controls on routine activities.

Such approach has been conducted in NPP's maintenance activities with satisfactory outcome. The instructions asked the subjects to think aloud what they were doing (Noizet, 1998). Findings show that routines are often evoked long or short before being implemented, with a clear and conscious effort of global activity planning, fixing tasks to be carried out with attention, and pre-activating check points of others. This result tends to advocate the idea that cognition is managing workload and risk much before task completion. Planning is not only a matter of task/goal ordering and decomposition, but also a matter of global strategy for efficient completion. The performance demand (stakes, level of requirements), the context (technical and social), and the metaknowledge (familiarity, fatigue) are the most frequent items associated to this planning activity of routines.

♦ *Interviews techniques and autoconfrontations*

Traditional interview techniques after or during work activity are of various interest for the study of routines:

1/ "Situation Awareness" interviews interrupt the operator's activity at specific moments in order to empty the content of his work memory and determine the state of his representation of the situation. Such interruptions have been tested and are useless for the study of routines because of the immediate rationalization.

2/ 'Auto-confrontation' have been carried out with maintenance crewmembers on various nuclear EDF sites (Noizet, 1998). The data was transcribed using a predicate/argument coding scheme (see Hoc & Amalberti, 1999). The findings are relatively poor for the purpose of modeling routine activities. Again, the operators tend to rationalize and skip the very nature of the routine activities to the benefit of conscious activities.

3/ A third strategy consists in using paper and pencil protocols, asking operator to comment on. Comparison between operators is possible, but the absence of the pace of real activity have proved these interviews unsuited for the study of routines.

4/ Last, semi-guided interviews and/or questionnaires have been used for subjects of more common interest, not directly linked to the activity analyzed. They are relevant to the study of the operators' perception of routines and of the structure of the organization in which they take place.

More "targeted" methods

Diary:

Participants in the *diary* must write down how they managed events defined beforehand by the researcher. According to Sellen (1994), who has used this method to study the mechanisms of detection and recovery of routine errors, this method can provide clues to non-conscious processes. However, though a detected error is an event sufficiently relevant to be reported (surprise effect), a "successful" routine goes unnoticed (absence of conscious awareness of its success).

'Explicitation' Interviews

These techniques aim at encouraging, helping and provoking a descriptive report on how a task is actually carried out (Vermersch, 1993). Several techniques are used to encourage verbalization of the implicit and low level process of consciousness implicated in action (Vermersch talks about pre-reflexive cognitive know-how): setting up of a specific relational framework; tools for channeling the content of verbalization; use of behavioral signs of the explicitation process; use of a descriptive questioning mode (primacy of the "how" over the "why") associated with specific rebounding questions (e.g. in the case of denial) in order to facilitate the consciousness process.

These techniques have been partially implemented in interviews with NPP operating crewmembers confronted with the film of their activity while performing a very familiar task (Noizet, 1998). This method has yielded valuable data for the modelization of the cognitive control of routines: fluctuations in the control level; specification beforehand of the elements of a situation which will be given attentional control; phenomenon of diffused supervision ensuring identification of shifts whatever the control level of the activity, and the expression of routine information taking. However, the delicate implementation of this method in the field (introduction of a camera, depth and length of interviews), its dependency on the limitations of observation (no routine has been categorically identified), and the arduous analysis of this specific verbal data, have contributed to diminishing its achievements.

An alternative approach: Simulation of Work

To sum up, most of the methods mentioned before do not overcome satisfactorily the difficulties encountered in reaching routine actions. The ideal solution would be to define reliable behavioral indicators on which the collection of verbalization could rely on, but without provoking a higher control level of the operators. The simulation of familiar tasks in a recreated work environment represents a promising path to go that way. It creates a paradigmatic situation sufficiently remote from the real situation so as not to interact with the activity naturally implemented; however, it remains close enough to reality to allow operators transferring a significant part of their routines when carrying out the simulated task.

From the simulation of field interventions to experimentation

The electrical building of units 1 - 2 of the Tricastin power plant (Drôme - France) have been recreated virtually through the computer editing of photographs allowing for movement and interaction. Selected on the advice of experts, fourteen field interventions are simulated taking into account their operation context and expected feedback. Movement and manipulations are done through the use of contextual mouse pointers which symbolize the action to be carried out at the place indicated on the image. This simulation gives field operators total freedom in their choice of a path and management of simulated interventions.

Each photograph is characterized by a name, a type of function (Movement, Interaction with the system, Door), and the number of actions on the view. The device enters the following data: the sequence of images chosen to carry out the intervention and the Time Spent per View (TSV).

Assuming that any cognitive activity requires a minimum of cognitive resources and time (Navon & Gopher, 1979) the hypothesis is that the views characterized by a significantly high TSV are views which generate a specific cognitive demand or a change in cognitive control. Interviews carried out by confronting field operators with these significantly different views allow to explicit the object, the reason, and the result of this higher involvement of cognition. Thus, this simulation offers a unity of analysis of expert operators' activity: the view; and a means for following the evolution of the cognitive system's involvement in the progress of the routine activity: Time Spent per View (TSV).

This simulation is used in the framework of experimental protocols to test the authors' control model of routines.

Following a phase of familiarization with the device, expert field operators are asked to perform a familiar intervention as they perform it in real life. This intervention is repeated until its performance reaches a stable form (sequence of views and TSV identical from one test to the other). Modifications of the performance context (double task, manipulation of the environment) are then introduced and their impact measured in relation to the stabilized form of performance of the familiar intervention. In a post-experimental interview the field operators are confronted with the views which stand out significantly from the others in terms of TSV, with the errors made and recovered, and with the most significant shifts induced by the experimental manipulations on the stable form of performance of the intervention. This quantitative and qualitative data is used in the framework of an analysis of the participants' activity focused on the cognitive control of routines involved in simulation.

First results

To date, the only available data comes from a phase of validation and pre-experimentation involving two field operators. According to this data, operators use their natural expertise to carry out simulated familiar interventions. They take the usual path and operating mode, and some results suggest they transfer to simulation the routines performed in real life.

1/ The data obtained suggests no learning process. The participants choose from the outset a path and a rhythm of "movement" of their own which they reproduce from one test to the other.

2/ The operators have difficulty in performing a less familiar task. Simulation does not provide enough information to enable them to validate their hypotheses regarding the localization of equipment.

3/ Operators are at a complete loss when a wrong clic sends them suddenly off the routine path. This situation is difficult to recover since it does not correspond to any reality.

The validation phase has shown that simulation is valid for carrying out routine activities, and valid, therefore, for studying the cognitive control of routines. This is confirmed by a brief data analysis. This data reveals the existence of points of control, which differ from one operator to the next, thus highlighting the role of individual experience in the cognitive management of routines. Besides, modification of the context of intervention has an impact on the way in which it is carried out (data obtained during a test performed on an operator and comprising 4 modifications in the conditions of the context - time pressure and double task). One observes the appearance of errors and a mean reduction of the TSV, which differs according to the type of view and experimental conditions. This time saving is done on the views usually characterized by a low TSV (delegation to routines), and certain points of control disappear whereas others remain invariant (optimization of the control). More results will be available at the conference presentation.

REFERENCES

Amalberti, R. (1996), La conduite des systèmes à risques. Paris: PUF

Anderson, J.R. (1982), Acquisition of cognitive skills. Psychological Review, 89, 369-406.

Baars, B.J. (1992), The many uses of error : Twelve steps to a unified framework. In B.J. Baars (Ed.), Experimental slips and human error : Exploring the architecture of volition. NY : Plenum Press, pp. 3-34.

Baars, B.J. (1997). In the theatre of consciousness : Global Workspace Theory, a rigorous scientific theory of consciousness. Journal of Consciousness Studies, 4(4), pp. 292-309.

Dubuisson, S. (1998). Regard d'un sociologue sur la notion de routine dans la théorie évolutionniste. Sociologie du Travail, 4, 491-502.

Hoc, J.M. & Amalberti, R. (1999). Analyse des activités cognitives en situation dynamique: d'un cadre à une méthode. Le Travail Humain, 62, 2, 97-129

Kellog, R.T. (1982), When can we introspect accurately about mental processes ? Memory and Cognition, 10, 141-144.

Logan, G.D. (1988). Toward an instance theory of automatization. Psychological Review, 95, 492-527.

Navon, D. & Gopher, D. (1979). On the economy of the human processing system. Psychological Review, 86, 214-255.

Noizet, A. (1998). Compréhension et modélisation de la gestion des activités routinières des agents de terrain en centrale nucléaire. Rapport LAMIH, Convention EDF/NEB n°T54/B28 865/0/RNE695. Valenciennes..

Norman, D.A. & Shallice, T. (1986) Attention to action : Willed and automatic control of behavior. In R. Davidson, G. Schwartz, & D. Shapiro, Consciousness and Self-regulation, Vol.4 (pp.1-18), NY: Plenum Press.

Rasmussen, J. (1986). Information processing and Human-Machine Interaction : An approach to cognitive engineering. Amsterdam : North Holland.

Reason, J. (1984). Absent-mindedness and cognitive control. In J.E. Harris & P.E. Morris (Eds.), Everyday memory, actions and absent-mindedness. London : Academic Press, pp.112-132.

Reynaud, B. (1998). Les propriétés des routines : Outils pragmatiques de décision et modes de coordination collective. Sociologie du Travail, n°4/98, 465-477.

Rizzo, A., Bagnara, S., & Viscolia, M. (1987). Human error detection processes. International Journal of Man-Machine Studies, 27, 555-570.

Sellen, A.J. (1994). Detection of everyday errors. Applied Psychology, 43 (4), 475-498.

Shiffrin, R.M. & Schneider, W. (1977). Controled and automatic human information Processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 127-190.

Vermersch, P. (1993), Pensée privée et représentation dans l'action. In A. Weil, P. Rabardel, & D. Dubois (Eds.), Représentation pour l'action. Toulouse : Octarés, pp. 209-232.

RÉSUMÉ

PROCESSUS MENTAUX ROUTINIERS DANS LES ACTIVITÉS SUR LES INSTALLATIONS DES CENTRALES NUCLÉAIRES : ASPECTS LIÉS À LA TRAÇABILITÉ ET À LA FIABILITÉ

L'étude répond à une demande d'EDF pour réduire l'occurrence des erreurs de routine de type confusions chez les opérateurs. Le texte se limite à deux aspects: (i) la discussion de la définition des routines (terme assez polysémique et flou), et (ii) l'analyse critique des méthodes permettant d'accéder à ces routines. Différentes techniques d'entretiens, de questionnaires, et une méthode plus expérimentale basée sur une simulation virtuelle sont testées. Au delà de ces méthodes, l'étude ambitionne de comprendre quel est le rôle des routines dans le réglage dynamique de la cognition et la gestion des risques.

MOTS CLÉS : Routines, Contrôle cognitif, Compromis cognitif, Méthodologie, Simulation, Centrales nucléaires

Session 6: Error management
La gestion de l'erreur

DISTURBANCE MANAGEMENT OF COMPLEX DYNAMIC SYSTEMS USING MFM

Johannes Petersen

Department of Automation, Technical University of Denmark
DK-2800 Lyngby, Denmark, jop@iau.dtu.dk

ABSTRACT

When managing familiar disturbance situations in complex plants human operators utilize their previous experience to cope with system constraints in a cognitive economical way. In unfamiliar disturbance situations, however, judgements based on previous experiences alone are error prone and there is need for decision support systems. This paper discusses some important aspects of such systems, emphasizing the phase of problem formulation. Problem formulation is particular important for successful disturbance management and it is illustrated, by means of an example, how Multilevel Flow Modeling (MFM) could provide support for this task.

Keywords

Disturbance management, diagnosis, situation assessment, problem formulation, MFM, process control.

INTRODUCTION

When managing disturbance situations in complex plants human operators have to cope with different kinds of complexity measures, such as the *causal structure*, i.e. the physical interactions between subsystems, as well as the *goal structure*, i.e. the strong and often conflicting demands on aspects as economy, production and safety (Lind, 1996). It seems that operators learn to cope effectively with such constraints in a wide range of situations. One explanation could be that operators manage incidents by means of recognizing a situation as typical, making judgements of prototypicality (Rosch, 1975). This is the underlying assumption of the so-called *recognition-primed decision model*, proposed by Klein and his colleagues (e.g. (Klein, 1989), (Klein, 1991)). This model might explain how decision makers (e.g. fireground commanders and tank platoon leaders (Klein, 1989)), in their natural environments, manage to cope with incidents in a cognitively economical way, without considering action alternatives in the concrete situation. However, in complex plant supervision where operators have to cope with less familiar situations it is a problematic strategy. Managing such situations requires knowledge based reasoning, i.e. reasoning on the basis of models of the system being supervised.

This claim is supported by a recent study of nuclear power plant operators performance in cognitively complex simulated emergencies, made by Roth (Roth, 1997). In (Roth, 1997) it is concluded: "In this study we found little evidence of pure recognition-primed decision making....Operator decision making was based on active construction of the factors known or hypothesized to be influencing plant state at any given point in time...In our study we found evidence of situations where crews needed to utilize mental models of physical plant systems and to reason qualitatively about expected effects of different factors influencing plant state in order to localize plant faults and identify actions to mitigate them. There is need to foster and support accurate mental models through training and control room display and decision aids.... Another type of needed knowledge concerns important plant goals and means to achieve them. Our study found evidence that operators needed to reason about plant goals, and evaluate alternative means to achieve them"

In the present paper we discuss some important aspects of knowledge based support systems for disturbance management of complex dynamic systems. We will concentrate particularly on unfamiliar (non-routine) situations where there is special need for constructive problem formulation based on reasoning about means and ends. The operator has to formulate high level goals and strategies based on an overall assessment of the current situation and possible means of interventions. It is believed that problem formulation is particular important for successful management of disturbances. Furthermore, it is a demanding task and knowledge based support is needed.

PROBLEM FORMULATION IN DISTURBANCE MANAGEMENT

In unfamiliar disturbance situations or emergencies operators will have difficulties to understand the situations and hence what to do. He is faced with a *problematic situation*, perceiving a discrepancy between the actual behavior of the system and the 'normal' behavior. The initial phases of disturbance management resembles an ill-structured task where neither the *goals* nor the *frame of reference* are well-defined (Simon, 1995). Rasmussen points to the circularity of such situations (Rasmussen, 1986): "The action to take depends on the result of the diagnosis. The diagnostic process, however, depends on the goal to pursue which, in turn, depend on the result of a diagnostic judgement." It is not meaningful to limit the scope of diagnosis to localization of faults we have to see diagnosis in a broader perspective as an integrated part of decision making intimately connected with action (Rasmussen, 1986), (Rasmussen, 1993).

Problem Framing and Situation Assessment

In order to handle problematic situations it must be cast in a conceptual space (model) which frames the problem and provides rules for reasoning about it. Schön refers to this process as *problem framing*: "When we set the problem we select what we will treat as the "things" of the situation, we set the boundaries of our attention to it, and we impose upon it a coherence which allows us to say what is wrong and in what directions the situation needs to be changed." (Schön, 1983), "...it is through naming and framing that technical problem solving becomes possible." (Schön, 1987)

It is clear that a proper problem framing has to be done on the basis of some background knowledge. Of course, the operator bring background knowledge to the situation implicitly via the *activity* with which he is currently engaged (Clancey, 1997). Being engaged with an activity provides certain expectations and anticipations of the situation. In the initial phase of an unfamiliar situation, however, such implicit structures typically are inadequate and the operator has to make an active framing of the problem from a systemic perspective. The reason for this is that the task of the operator is subordinate to the *goal structure* of the system being supervised. The goal structure of the system provides important framing values, which the operator is supposed to take into account when framing problems in unfamiliar disturbance situations. Typically, the goal structure is complex, comprising different types of goals, achieved in different ways. For instance, *production goals* prescribe what has to be achieved (in normal operation), whereas *safety goals* prescribe something that has to be prevented.

A proper problem framing should provide a conceptual space in which it is possible to evaluate the current situation as well as its possible future implications in relation to important plant goals, at the same time, taking into account possible means of interventions. We will refer to this type of high level reasoning as *situation assessment*. Situation assessment provides the basis for initial formulation of *goals* and *strategies*, directing subsequent problem solving activities. Situation assessment requires demanding on-line reasoning and there is need for systems that support this kind of task. Unaided judgements can be insufficient and even critical. The operator might reason on the basis of inadequate models of the system or simply resort to lower levels of decision making (e.g. recognition-primed decisions). From the point of view of *problem formulation* erroneous assessment of the situation might lead to errors of solving the wrong problem (Woods, 1988).

The initial task in an unfamiliar disturbance situation will likely be to *compensate* the current or future influence on vital performance parameters (Rasmussen, 1993). Later we shall indicate how MFM models can support reasoning activities during situation assessment providing a useful systemic problem framing. In the following section we will outline important factors of problem solving.

The Role of Goals, Strategies and Expectancies in Dynamic Problem Solving

In the phase of problem formulation high level goals and strategies are formulated. Such constructs are of course provisional but have the important function of guiding the subsequent problem solving activities, aiming at implementation of the strategies through more concrete courses of actions. The goal directs the operator's attention towards information significant for solving the specific problem (compare with John Dewey's theory of reflective action (see (Miettinen, *in press*) for a summary) and the *competent* level in Dreyfus and Dreyfus' stage model of expertise (Dreyfus, 1986)). Apart from being necessary for problem solving a goal-directed perspective is of course also critical as it biases the operator's attention to certain local aspects of the situation preventing him from perceiving the situation at a more global level. Relying too persistently on the goals formulated, in a dynamic environment, may lead to fixation errors, i.e. failure to revise situation assessment (De Keyser, 1990).

The operator can utilize prognosis of system evolution made during situation assessment to judge if the system evolves as expected. If it does not, the operator should realize that the problem formulated might no longer be valid, since the unexpected evolutions were not taken into account at the time of problem formulation. In that way operators expectancies can be utilized to judge the validity of a formulated problem. In situations where expectancies are violated, revisions should be made in order to avoid fixation errors. It is important that such revisions, if necessary, are done on the basis of a global reassessment of the situation. Of course there might be other reasons for a needed reassessment of the situations than violated expectancies, e.g. unexpected problems might occur when trying to implement strategies through concrete actions.

MULTILEVEL FLOW MODELING

Multilevel Flow Modeling (MFM) is developed by Lind (see, (Lind, 1990), (Lind, 1994) and (Lind, *submitted for publication*). See also (Petersen, 1998) for some recently introduced relations. MFM represents intentional knowledge in the domain of complex continuous process plants for the purpose of diagnosis, planning and interface design. It employs abstractions along *means-end* and *part-whole* dimensions and represents multiple functional levels of a process plant in terms of mass and energy flows. These levels are depicted as so-called *flow structures* comprising networks of connected mass and energy flow functions. Functions are abstract interpretations of process behavior that is considered useful in view of goals. Flow structures achieve objectives. *Objectives* are concrete technical specifications of how *goals* (expressing system requirements or human needs) are implemented according to a designer's choice. Note that a goal can often be met through the achievement of several alternative objectives. By means of explicit relations between goals, objectives and flow structures the individual functional levels are linked in an overall network that capture the intentional structures of the system at different levels of detail. By means of a small example we will explain some aspects of the MFM methodology, but in the following, we take it that the reader is acquainted with the basic principles of MFM.

A Simple Example

The example is a simplified furnace plant shown in Figure 1, adapted from (Gofuku, 1997). The ore is supplied to the furnace, where it is melted by means of the heat from the burner. It is discharged from the furnace as melted metal. The MFM model of this plant is shown in Figure 2 (note that the model is different than the one given in (Gofuku, 1997)). The labels in Italics are annotations facilitating the reading of the MFM model.

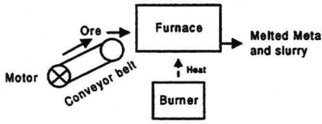


Figure 1

The top objective O1 states that the production of melted metal is maintained. This objective is achieved by a mass flow structure FS1. FS1 represent the flow of ore into the furnace and the discharge of melted metal. The supply of ore is represented by the transport function F2 and the discharge flow is represented by the transport function F4. F2 is a *product* of the process "giving work to ore" and a PP (producer-product) relation between FS2 and F2 represents this. FS2 represents the flow of momentum from the motor to the conveyor belt. In FS1 the transport function F4 is conditioned by the objective O2, stating that the melting of ore is maintained. The achievement of O2 *enables* the discharge flow from the furnace. The objective O2 is achieved by the energy flow structure FS3, representing the heat balance in the furnace. Heat goes into the furnace from the burner and leaves the furnace via the furnace wall (F17) and via the ore leaving the furnace as melted metal. The heat leaving the furnace via the ore is represented by a transport functions (F14). This function is *mediated* by the supply of ore, represented by a M (mediate) relation from FS1 to F14. A safety objective O3, stating that a too high temperature in the furnace should be prevented, is achieved by the flow structure FS3. For reasons of simplicity the *goals* are left out in the MFM model, only the *objectives* are shown.

The PP (producer-product) relation and M (mediate) relation are proposed as conceptual clarification of the interrelations between flow structures in MFM (see, (Petersen, 1998)). The PP relation in the example prescribes a causal relation between the flow of momentum in FS2 and the supply of ore (F2) in FS1. The M relation represents a causal relation between the supply of ore in FS1 (F2) and the removal of heat from the furnace via the melted metal (F14). *Points of intervention* indicate the functional association of properties that can be manipulated by the operator (Lind, *submitted for publication*).

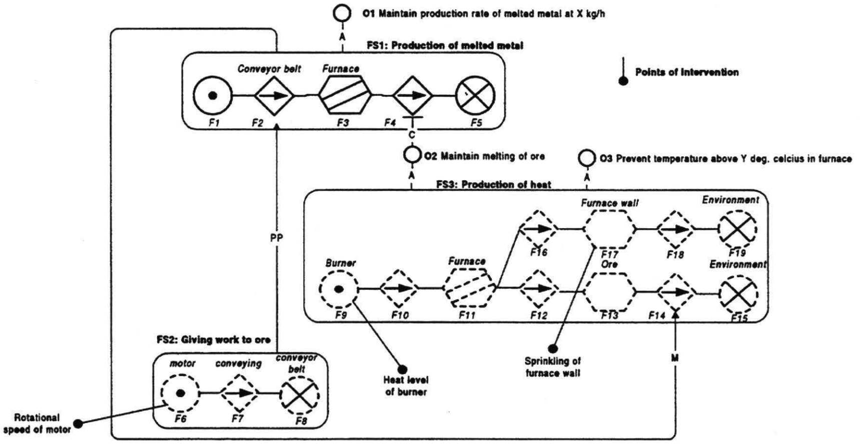


Figure 2

SITUATION ASSESSMENT USING MFM

MFM provide a representation of a system that integrates the goal structure and the causal structure by means of a functional representation of the mass and energy flow structures of a system. The functional structures achieves objectives (and hence goals) and are realized through system behavior. By means of rules of interpretation sensor measurements can express on-line information about the state of functions and goals (see (Larsson, 1996)). That is, MFM provides a framing in terms of the goal function structure of the system from the point of view of the designer. In that sense it is an objective framing that will support the operator in unfamiliar situations to perform an assessment of the situation on the systemic level. Due to the functional categorization MFM provide a useful framing also in unanticipated disturbance. The reasoning possibilities and the explicit indication of intervention points and their possible functional association enable a formulation of high level strategy of intervention. It is important to integrate reasoning about goals and strategies in the same framework – MFM makes that possible.

In summary, MFM provides an instrumental framing as well as a reasoning aid for constructing *causal hypotheses* (diagnosis and prognosis) in relation to possibilities of intervention. Such capabilities would support operators to make an assessment of the situation and formulate a provisional problem, i.e. goals and strategies. The goals and strategies serve to initiate and guide problem solving activities. In the following we will consider a simple example of how MFM could support situation assessment. In the example we draw on the results from (Gofuku, 1997) and (Petersen, 1998).

Example of Situation Assessment Using MFM

Suppose we have a disturbance situation, which in the MFM model, show as a disturbance of function F2 – the supply of ore is degraded. Generally, it is not possible, based on the information of the MFM model alone, to identify the exact cause, but only its functional consequence in a specific flow structure. In disturbance situations knowledge of functional consequences in relation to mass and energy flow structures is essential since the preliminary goal is not to *correct* the failure but to *compensate* the disturbance caused by it (Rasmussen, 1986).

Underlying the degraded transport function F2, representing ore supply, there might be several *causes*, e.g. "no ore is present", "the conveyor belt is stuck", "motor damage" etc.. From the MFM model the possible future *consequences* can be inferred. We will not go into details about the rules for reasoning, but since the mass flow structure FS1 is disturbed the *mediate relation* tells us that the heat flow structure FS3 will be affected. More specifically the transport function (F14) will degrade, and cause an imbalance in the heat flow structure FS3, which means that the temperature in the furnace is likely to rise if nothing is done to prevent this.

There are two goals associated with the heat flow structure FS3, a *production goal* and a *safety goal*. The safety goal states that a rise in the temperature of the furnace above a certain level should be prevented. Obviously, it makes little sense to ensure the production goal as the mass flow structure it support is disturbed anyway because of the disturbed transport function F2 (of course there might be reasons to do it anyway, but we will not consider them here). It seems that the most appropriate thing to do is to ensure that the safety goal will not be violated. In order to do that we must compensate the possible disturbance of the heat balance of the furnace, either decreasing the inflow or increasing the outflow of heat. Hence, two things can be done: 1) reducing or shutting down heat production of the burner or 2) sprinkling the furnace wall. The actual choice between these alternatives will depend on further judgements by the operator. If the cause can be recovered relatively fast it might not be needed to shut down the burner but only to reduce heat production. It might even be sufficient to sprinkle the furnace wall to prevent an increase in the temperature in the furnace. It is believed that the qualitative nature of the reasoning based on an MFM is a desired feature of decision support systems (see the above quoted conclusions in (Roth, 1997)), which have to be used under great time pressure.

DISCUSSION

This paper has emphasized the importance of problem formulation in unfamiliar disturbance situations of complex dynamic plant systems. It is crucial that the initial formulation of the problem is based on a proper assessment of the situation in relation to important goals and functions of the system being supervised. In this paper we have indicated, through a simple example, how MFM models could provide useful for the framing of unfamiliar situations from a systemic point of view. MFM represents the goal function structure of the system, which, especially in the initial phases of a disturbance management, represents important background knowledge for the process of framing the problem.

Apart from representing knowledge that is useful for problem framing purposes MFM enables reasoning about possible *causes* and *consequences* of the situation in relation to system goals and function, as well as the possible *interventions*. This type of framework for high level reasoning based on multiple levels of abstraction could support the operator in the phase of problem formulation enabling him to perform an overall assessment of the situation and formulate high level goals and strategies. During the process of situation assessment the operator forms expectancies which will help him to judge the future evolution of the system as well as the effect of his interventions. Both an initial problem formulation and expectations are essential for the prevention of fixation errors.

Although the shown example is simple it is believed that it illustrates the basic principles of how MFM could provide useful as an underlying basis for representation aids or decision support system. However, as mentioned by Lind (Lind, *submitted for publication*), MFM does not address how to present or visualize information.

Goals and strategies are provisional and the problem solver may need to reassess the situation during problem solving activities. Although, MFM provides a useful framework for reasoning activities during situation assessment in the initial phase of disturbance management, this is not evidently the case when the operator has to reassess the situation later in the problem solving process. The appropriate framing depend on the state of the problem solving process (refer to the context-sensitivity problem defined (Woods, 1995)). That is, although the intentional structure represented in an MFM model is more or less invariant a framing based on this knowledge is not useful in all phases of problem solving.

REFERENCES

Clancey, W.J. Knowledge, Situations, and Activity, in: *Expertise in Context*, Feltovich, P.J., Ford, K.M., Hoffman, R.R.(eds.). AAAI Press/ MIT Press, 1997
De Keyser, V. & Woods, D.D. (1990). Fixation Errors: Failures to Revise Situation Assessment in Dynamic and Risky Systems, in: *Systems Reliability Assessment*, 231-251, Colombo, A.G. and de Bustamante, A.S. (eds.)
Dreyfus, H.L. & Dreyfus, S.E. (1986) Mind over Machine : The Power of Human Intuition and Expertise in the Era of the Computer. Free Press.

Gofuku, A. & Tanaka, Y. (1997). Development of an Operator Advisory System: Finding Possible Counter Actions in Anomalous Situations. *Proceedings of the Fifth International Workshop on Functional Modeling of Complex Technical Systems*, Paris-Troyes, France, July 1-3, 1997

Hoc, J.-M., Amalberti, R. & Boreham, N. (1995). Human Operator Expertise in Diagnosis, Decision-Making, and Time Management, in: *Expertise and Technology*, Hoc, J.-M., Cacciabue, P.C., Hollnagel, E.(eds.), Lawrence Erlbaum Associates, Publishers.

Klein, G.A. (1989). Recognition-Primed Decisions. *Advances in Man-Machine Systems Research, 5*, 47-92.

Klein, G.A. & Calderwood, R. (1991). Decision models: Some Lessons from the field. *IEEE Transactions on System, Man, and Cybernetics, 21*, 1018-1026.

Larsson, J.E. (1996) Diagnosis Based on Explicit Means-End Models. *Artificial Intelligence, 80*, 29-93.

Lind, M. (1990). Representing Goals and Functions of Complex Systems, Tech. Rep. 90-D-381, Department of Automation, Lyngby, Denmark

Lind, M. (1994). Modelling Goals and Functions of Complex Industrial Plants. *Applied Artificial Intelligence, 8*, 259-283.

Lind, M. (1996). Interpretation Problems in Modelling Complex Artifacts for Diagnosis. *Proc. Cognitive Systems Engineering for Process Control, CSEPC'96, Nov. 12-15, Kyoto, Japan.*

Lind, M. (*submitted for publication*). Plant Modeling for Human Supervisory Control.

Miettinen. R. (in press). The concept of experimental learning and John Dewey's theory of reflective thought and action., *International Journal of Life Long Education*, 19 (1), 2000

Petersen, J. & Lind, M. (1998). Considerations on Agents and Objects in Multilevel Flow Modelling. *Proc. European Annual Conference on Human Decision Making and Manual Control, Dec. 14-16, Valenciennes, France.*

Rasmussen, J. (1986). Information Processing and Human-Machine Interaction. An Approach to Cognitive Engineering, *Amsterdam: North-Holland.*

Rasmussen, J. (1993). Diagnostic Reasoning in Action. *IEEE Transactions on Systems, Man, and Cybernetics, Vol 23 (4)*, 981-992.

Rosch, E. & Mervis, C.B. (1975). Family Resemblance: Studies in the internal structure of categories. *Journal of Cognitive Psychology, 7*, 573-605.

Roth, E. (1997). Analysis of Decision Making in Nuclear Power Plant Emergencies: An investigation of Aided Decision Making, in: *Naturalistic Decision Making*, Zsambok, C.E. and Klein, G. (eds.). Lawrence Erlbaum Associates, Publishers.

Schön, D. A. (1983). The Reflective Practitioner. Basic Books

Schön, D. A. (1987). Educating the Reflective Practitioner. Jossey-Bass Publishers

Simon, H.A. (1995). Machine as Mind, in: *Android Epistemology*, Ford, K.M., Glymour, C., and Hayes, P.J.(eds.). AAAI Press/ MIT Press.

Woods, D.D. (1988). Coping With Complexity: The Psychology of Human Behaviour in Complex Systems, in: *Tasks, Errors and Mental Models, Goodstein*, L.P., Andersen, H.B. and Olsen, S.E.(eds.), Taylor and Francis.

Woods. D.D. (1995). Toward a Theoretical Base for Representation Design in the Computer Medium: Ecological Perception and Aiding Human Cognition, in: *Global Perspectives on the Ecology of Human-Machine Systems,1*, Flach, J., Hancock, P., Caird, J. and Vicente, K. (eds.). Lawrence Erlbaum Associates Publishers.

RÉSUMÉ

L'UTILISATION DE MFM POUR LA GESTION DES INCIDENTS DANS LES SYSTÈMES DYNAMIQUES COMPLEXES

En gérant des situations incidentelles familières sur des installations complexes, les opérateurs humains utilisent leur expérience antérieure pour traiter les contraintes du système avec un coût cognitif réduit. Cependant, dans les situations incidentelles non familières, les jugements uniquement fondés sur les expériences antérieures peuvent conduire à des erreurs et il faut des systèmes d'aide à la décision. Cette communication discute de quelques aspects importants de ces systèmes, en insistant sur la phase de formulation du problème. Cette phase est particulièrement importante pour la réussite dans la gestion de l'incident et un exemple illustre la façon dont la modélisation multi-flux (MFM) peut apporter un soutien dans cette tâche.

MOTS CLÉS : Gestion d'incidents, Diagnostic, Évaluation de situation, Formulation du problème, MFM, Contrôle de processus.

ERROR RECOVERY IN SOCIO-TECHNICAL SYSTEMS

T.W. van der Schaaf
Safety Management Group, Eindhoven University of Technology
P.O. Box 513, Pav. U-8, 5600 MB Eindhoven, The Netherlands
tel. +31 40 2474380; e-mail: T.W.v.d.Schaaf@TM.TUE.NL

L. Kanse
Human Reliability Associates
HRA-NL: Arenberglaan 280, 4822 ZR Breda, The Netherlands
tel. +31 76 5430028; e-mail: superbear@compuserve.com

Abstract: This paper highlights the positive role that human operators often play in preventing small failures and errors from developing into an actual system breakdown. The resulting 'near misses' may provide an insight into a powerful alternative to human error prevention, namely: human recovery promotion. Theoretical approaches to modeling error recovery are discussed and translated into empirical research questions. These are partly answered by a number of pilot studies. The main conclusions are that error recovery is much more than simple luck or coincidence, that its root causes can be identified, and that these should have design implications for the technical and organisational context of the human operator's task.

Keywords: error recovery, human reliability, incident analysis, near misses

INTRODUCTION

The basic focus of safety management so far has been on the prevention of errors and failures. When taking a closer look at what exactly we want to prevent, we find that it is rather the (negative) consequences of a failure than the occurrence of the failure itself (Frese, 1991; Frese et al., 1996; Kanse & Embrey, 1998). This idea leads to a relatively new area of research, namely that of recovery promotion. Recovery factors are those factors that contribute to (complete or partial) recovery once an error or failure has occurred, thus preventing or reducing the negative consequences of that error or failure. This paper describes a research project focussing on the positive role that human operators often play in the recovery process.

In the next section first a simple incident causation model is presented in which the presence or absence of successful human recovery plays a decisive role in determining the effects of process deviations, technical failures and errors on the safety and reliability of socio-technical systems. Then, the process of human recovery is described to consist of three phases: detection of symptoms; localisation of their cause(s); and correction to return the system to its normal status. A discussion follows of the different recovery patterns that can be distinguished after an error is detected. The following sub-section then deals with the relationship of human error causes and the probability of recovery, and with the error detection lag. These theoretical predictions are mainly based on well known cognitive limitations and feedback-related aspects of the task situation.

Then four ways of classifying (human) recovery in actual process control situations are proposed. The first classification deals with the type of preceding failure(s), for instance technical, organisational or human failure. The second classification locates the phase in which a recovery factor contributes to the recovery process: detection, localisation, or correction. Another way to look at human recovery is to distinguish the reaction after symptom detection: ignore the deviating status; repeat sequence of actions; or attempt fault localisation and correction. The fourth and most important possibility is to categorise the factors in the socio-technical system that triggered or enabled recovery: technical factors related to process or interface design; the organisational and management context; and operator factors.

In the third section, these theoretical approaches and existing insights are translated into specific empirical research questions on which the described research project focussed. A set of recent pilot studies of incidents in a steel plant and an energy production unit, as well as medical errors in a surgical ward, will then be presented and their results will be used to formulate initial answers to the research questions. Data from a petro-chemical plant is included in the fifth section, which shows again the importance of human operators in the recovery process and the difference between factors contributing to failure and those contributing to recovery. Finally implications of the research findings for designing recovery into socio-technical systems will be discussed.

THEORETICAL APPROACHES & EXISTING INSIGHTS

Incident Causation Model

In Van Vuuren (1998) a simple incident causation model is used (see figure 1) to define accidents, near misses and their common root causes consisting of technical, organisational and human (operator) factors. When incident development cannot be stopped by the system's predetermined barriers (Svenson, 1991) and lines of defence, the only distinguishing factor between an accident and a near miss is the presence or absence of successful 'accidental' or unplanned recovery.

Although actual accidents also may contain attempts at recovery, it is obvious that near misses as defined above are the optimal source of data to study the phenomenon of recovery as the positive counterpart of failure. The Eindhoven Classification Model ECM of system failure (Van der Schaaf, 1991 & 1992, Van Vuuren, 1998) has succeeded so far in modeling the failure factors (or root causes).

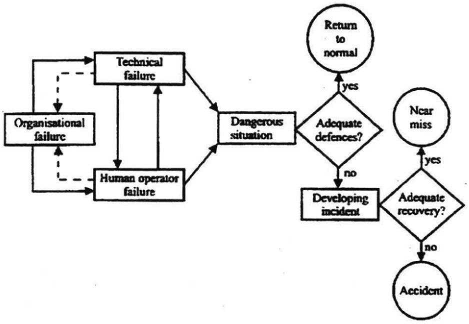


Figure 1. The incident causation model

Human Recovery Process Phases

The following definition of human recovery is proposed by Van der Schaaf (1988): "the (unique?) feature of the human system-component to detect, localize and correct earlier component failures. These component failures may be either his own previous errors (or those of colleagues) or failing technical components (hardware and software)".

This definition implies the following phases in the recovery process:

- detection: of deviations, symptoms, etc.
- localisaton: of their cause(s) (diagnosis in the strictest sense)
- correction: of these deviations by timely, effective counter actions, after which these deviations are nullified and the system returns to a stable status again

The error handling process as described by Zapf and Reason (1994) distinguishes three main process phases: the error itself, error diagnosis and error recovery. The error diagnosis phase consists of error detection and error explanation (corresponding with Van der Schaaf's detection and localisation phases), and the error recovery phase (which corresponds with Van der Schaaf's correction phase) consists of a planning and an execution step. Van der Schaaf's three recovery process phases can also be found back in Kontogiannis (1999), who distinguishes the error detection, error explanation and error correction aspects of the error handling process.

Most of the research on the recovery process has focussed on the detection phase. Sellen (1994) has identified action-based detection, outcome-based detection, and detection through limiting function as detection modes for the detection of slips and mistakes, and reminding/memory retrieval as detection mode for lapses. Bagnara et al. (1988) use a distinction that corresponds with Sellen's first three detection modes. According to them, detection can be the result of three types of mismatch between operators'expectations and the available information about the performed or to be performed activities, respectively: inner feedback, external feedback and forcing function.

Recovery patterns

As noted by Reason (1990) in his GEMS model, people seldom go through the entire analytic process of fault diagnosis when confronted with a deviation. This was confirmed by Brinkman (1990) who collected verbal protocols during a fault finding task. He observed the following three reactions after his subjects detected an error in their reasoning process:

- Ignore the error and continue: rely on system redundancy and subsequent error recovery factors.
- Simply repeat the most recent sequence of actions: try again, without any attempts at fault localisation.

- Attempt fault localisation and optimise corrective actions: either by forward analysis (repeat the most recent action sequence and check every step) or backward analysis (trace back from symptom detection to previous actions, until the error is found).

Bagnara et al. (1988) observed the same and distinguish six behavioural patterns after an error is detected, based on the amount and type of analysis that takes place: Immediate correction, automatic (fast) causal analysis, conscious causal analysis (a hypothesis is formulated), explorative causal analysis (more hypotheses are tested), and overcoming of the mismatch if all else fails.

Dependency of Recovery on Preceding Errors

Embrey and Lucas (1988) discuss several factors affecting the probability of recovery from error, and the error detection lag. This relationship is highly relevant to understand the role of feedback in the recovery mechanism and their main points may be summarized as follows:

- Causes of skill-based slips and lapses are relatively unrelated to subsequent recovery factors, their human recovery probability is high, and the error detection lag will be small.
- For rule- and knowledge-based mistakes the opposite holds: their recovery factors depend on the same preceding failure factors; probability of recovery is small, and the error detection lag large.

Main reasons for these predictions given by Embrey and Lucas (1988) include feedback related aspects and cognitive limitations: the awareness of an error possibility and the visibility of its effects is high for slips and lapses, but low for mistakes; cognitive limitations (e.g. confirmation-, fixation- and groupthink biases) would be small for slips and lapses, but large for mistakes. For our paper the main implication is that the nature of the preceding human error(s) should be highly predictive of any subsequent recovery.

A computer programming experiment by Bagnara et al. (1988) also indicates that a dependency exists between type of error, the detection mechanism and recovery patterns: Slips were most often detected via external feedback, and recovered via immediate correction without causal analysis. Rule-based mistakes had no preferred detection mechanism, and most often conscious or (less often) automatic causal analysis was undertaken to recover. Knowledge-based mistakes were most often detected via forcing function, and the most used recovery strategy was to undertake an exploratory causal analysis or otherwise to try to get rid of the mismatch.

Classification of Human Recovery Aspects

The preceding subsections lead to the following four ways of classifying (human) recovery aspects:

Classification according to Recovery Process Phase

Recovery root causes can be categorised according to whether they contribute to error or failure detection, to error localisation or diagnosis of what went wrong, or to the actual correction of the problem.

Classification According to Operator Reaction After Symptom Detection

As mentioned earlier, different recovery patterns used by human operators can be distinguished, based on the amount and type of causal analysis they undertake after error detection and the solution that is chosen to overcome the problem. Recovery can be classified according which to these recovery patterns is followed.

Classification Based on Preceding Failure

Both the ECM (Van der Schaaf, 1992, and Van Vuuren, 1998) and Embrey and Lucas (1988) provide the rationale for this taxonomy. Technical, organisational and human root causes of failures may be linked with their subsequent recoveries. Additional subcategories might include: recovery from one's own error, of from a colleague's; technical failure of equipment outside the central control room, of the interfaces within the control room, of process control software, etc.

Classification According to Type of Recovery Factor (or Recovery Root Cause)

Such a classification should be the most important one for the designers of socio-technical systems. The ECM for failure root causes, with various subcategories for technical, organisational, and human-skill, -rule and -knowledge based factors, could serve as a basis for recovery root causes too, with the following extensions:

- Technical design of the process: aim at maximum *reversibility* of process reactions (Rasmussen, 1986) to allow for correction, and 'linear interactions' plus 'loose coupling' (Perrow, 1984) of process components; these may be achieved by *structural* characteristics (e.g. buffers, parallel streams, equipment redundancy) and by *dynamic* characteristics (e.g. speed of process reactions, response delays).
- Technical design of the man-machine interface: aim at maximum *observability* (Rasmussen, 1986) to allow early detection of deviations and their effects (e.g. transparency instead of alarm inflation).
- Organisational and management factors: especially an updated, clearly formulated and well-accepted set of operating procedures, and a positive safety culture must be mentioned here (see also Van Vuuren, 1998).

- Human operator factors: optimize the cognitive capabilities (e.g. accurate mental process model) of operators through selection and (simulator-)training, but also by supporting them with software tools to test hypotheses and avoid certain biases.

EMPIRICAL RESEARCH QUESTIONS

Based on the proposals in the previous section, the human recovery research project of the Eindhoven Safety Management Group focusses on the following empirical research questions:

1) Is recovery more than sheer luck or coincidence? If so, then recovery can be built into a socio-technical system and be managed.
2) Can recovery be classified with the same root causes as those used for failures? If so, what is the contribution of human recovery in relation to technical and organisational failure barriers? How large is the contribution of human recovery in a variety of task situations and over a variety of system effects?
3) Are recovery factors identical to failure factors in a given socio-technical system? If so, then preventing errors and promoting recovery would have to focus on the same socio-technical system aspects.
4) In which phase(s) of the recovery process do recovery factors mostly contribute to system performance: symptom detection, fault localisation, or correction?

PILOT STUDIES

Pilot studies investigating the recovery process have been carried out in a steel production plant (Mulder, 1994), a surgical ward (Van der Hoeff, 1995) and an energy production unit of a chemical plant (Zuijderwijk, 1995). A variety of system effects have been investigated: safety, reliability, and environmental effects of system breakdown.

Method

The critical incident technique,using confidential interviews, (Flanagan, 1954) has been used in all pilot studies to collect data on a set of recent near misses in each of the pilot studies.

Results

The results of the pilot studies are summarised in the following tables (see also Van der Schaaf, 1996).

	# of analysed near misses	# of failure factors	# of recovery factors
steel plant	25	154	34
surgical ward	17	95	22
energy production unit	23	86	56

For the analysis of both the failure and the recovery factors, the subcategories of the ECM have not been used, only the main groups of technical (T), organisational and management (O) and human operator (H) factors, plus unclassifiable (X).

Failure factors	T	O	H	X
steel plant	21%	33%	44%	2%
surgical ward	6%	37%	56%	1%
energy production unit	40%	23%	33%	4%

Recovery factors	T	O	H	X
steel plant	47%	21%	21%	11%
surgical ward	16%	20%	55%	9%
energy production unit	27%	5%	66%	2%

Based a range of only 2 to 11% unclassifiable (that is: luck or coincidence) causes of recovery, a positive answer to research question 1 can be given: around 90 % or more of all recovery factors are clearly technical, organisational or human in nature and therefore eventually manageable. Comparison of the percentages of human recovery root causes with those of human failure root causes shows that they both vary between the case studies, respectively 21 to 66% and 33 to 56%. The human component should therefore also be taken seriously in terms of recovery possibilities (question 2).

Zuijderwijk (1995) shows that the patterns of failure and recovery factors are clearly different. Rule- and skill-based factors dominate the operator failures, while human recovery also includes knowledge-based insights as very important. Similarly, 'material defects' are the most prominent technical failures, while 'design' covers all technical recovery factors (see question 3).

Finally, data from the steel plant (Mulder,1994) showing that 26% of the recovery factors contribute to the detection phase, 9% to the localisation phase, and 53% to the correction phase, provides an answer to question 4: Hardly any recovery process goes through the more analytic localisation phase. Again, this could be interpreted as a confirmation of Reason's GEMS model, but there is also the possibility of an explanation in terms of time stress. If recovery is present only in the *very last phase* of the accident production chain of events (as was the case in most of the steel plant near misses) there may simply not be time enough for a time-consuming diagnostic effort; detection and correction 'just-in-time' may be all one can do in such cases.

DATA FROM A PETRO-CHEMICAL PROCESS PLANT NEAR MISS MANAGEMENT SYSTEM

Figure 2 is derived from data from the Near Miss Management System (NMMS) database of a large, international petro-chemical plant in Rotterdam. This NMMS contains more than 3000 reports so far and has been operational since 1994. Now circa 800 reports per year are added to the system. For those incidents that are selected for a causal tree analysis and classification following the ECM (10-20%), the figure shows the distribution across the different categories of root causes – both for failure factors and recovery factors. Once more the difference in distribution of the failure factors from that of the recovery factors is shown, and the obvious importance of the role of the human operator in recovery.

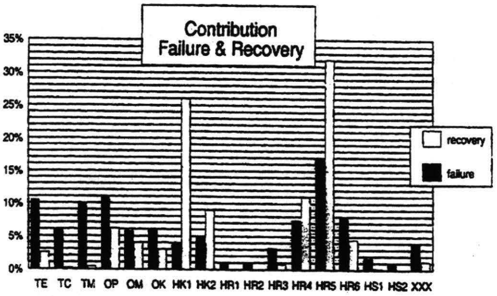


Figure 2. NMMS data from petro-chemical plant

IMPLICATONS FOR SOCIO-TECHNICAL SYSTEM DESIGN

In spite of the immaturity of the proposed models and classifications, and of the small number of recovery incidents gathered so far, these ideas and results are intriguing enough to formulate the following tentative implications for designing a socio-technical system:

- Consider recovery promotion as an alternative to failure prevention, especially when certain errors or failures are predictably unavoidable.
- Do not simply design out failure factors without considering the possible reduction of recovery factors: raising the level of automation in process control, or installing too many decision support tools for your operators may leave them helpless under certain situations.
- Try to support all recovery phases, detection (observability!), localisation and correction (reversibility!), primarily by means of an optimal man-machine interface.
- Invest in deep process knowledge of operators: reasoning beyond procedures appears to be essential for many recovery actions. Also,consider error management training (Frese, 1991): *learning to learn from errors* is perfectly in line with the concept of recovery promotion.

REFERENCES

Bagnara, S., Ferrante, D., Rizzo, A., & Stablum, F. (1988). *Causal analysis in error detection and recovery: when does it occur?* Paper presented at the international conference on joint design of technology, organization and people growth, Venice, October 12-14, 1988.

Brinkman, J.A. (1990). *The analysis of fault diagnosis tasks: Do verbal reports speak for themselves*? PhD thesis, Eindhoven University of Technology.

Embrey, D.E., & Lucas, D.A. (1988). The nature of recovery from error. In L.H.J. Goossens (Ed.), *Human recovery: Proceedings of the COST A1 Seminar on Risk Analysis and Human Error* . Delft: Delft University of Technology.

Flanagan, J.C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327-358.

Frese, M. (1991). Error management or error prevention: Two strategies to deal with errors in software design. In H.J. Bullinger (Ed.), *Human aspects in computing: Design and use of interacive systems and work with terminals* (pp. 776-782). Amsterdam: Elsevier Science Publishers.

Frese, M., van der Schaaf, T.W., & Heimbeck, D. (1996). Error Management and Recovery in Technical System Design. In: *Proceedings of the '96 Multiconference, Symposium on Robotics and Cybernetics,* July 1996, Lille, France, pp. 161-164.

Hoeff, N.W.S. van der (1995). *Risk management in a surgical ward (in Dutch).* MSc thesis, Eindhoven University of Technology.

Kanse, L., & Embrey, D. (1998). Recovery from failure. In: *Proceedings of 4th World Conference on Injury Prevention and Control,* May 17-20, Amsterdam, the Netherlands, p 1020.

Kontogiannis, T. (1999). User strategies in recovering from errors in man-machine systems. *Safety Science, 32,* 49-68.

Mulder, A.M. (1994). *Progress report 1993 of the SAFER project.* Report Hoogovens steel industry (in Dutch). IJmuiden: Hoogovens.

Perrow, C. (1984). *Normal accidents: living with high-risk technologies.* New York: Basic Books.

Rasmussen, J. (1986). *Information processing and human-machine interaction.* Amsterdam: Elsevier Science Publishing.

Reason, J.T. (1990). *Human Error.* Cambridge: Cambridge University Press.

Schaaf, T.W. van der (1988). Critical incidents and human recovery. In L.H.J. Goossens (Ed.): *Human recovery: Proceedings of the COST A1 Seminar on Risk Analysis and Human Error.* Delft: Delft University of Technology.

Schaaf, T.W. van der (1991). Development of a near miss management system at a chemical process plant. In: *Near miss reporting as a safety tool* (Schaaf, T.W. van der, D.A. Lucas & A.R. Hale (Eds.)). Butterworth Heinemann, Oxford.

Schaaf, T.W. van der (1992). *Near miss reporting in the chemical process industry.* PhD thesis, Eindhoven University of Technology.

Schaaf, T.W. van der (1996). Human recovery of errors in man-machine systems. In: *CCPS '96, International conference and workshop on process safety management and inherently safer processes,* October 8-11, Florida, USA. New York: American Institute of Chemical Engineers.

Sellen, A.J. (1994). Detection of Everyday Errors. *Applied Psychology: An International Review, 43*(4), 475-498.

Svenson, O. (1991). The Accident Evolution and Barrier function (AEB) model applied to incident analysis in the processing industries. *Risk Analysis, 11,* 499-507.

Vuuren, W. van (1998). *Organisational failure: An exploratory study in the steel industry and the medical domain.* PhD thesis, Eindhoven University of Technology.

Zapf, D., & Reason, J.T. (1994). Introduction: Human errors and error handling. *Applied Psychology: An International Review, 43*(4), 427-432.

Zuijderwijk, M. (1995). *Near miss reporting in reliability management.* MSc thesis, Eindhoven University of Technology.

RÉSUMÉ

LA RÉCUPÉRATION DES ERREURS DANS LES SYSTÈMES SOCIO-TECHNIQUES

Cette communication souligne le rôle positif que jouent souvent les opérateurs humains quand ils font en sorte que les petites défaillances et les erreurs minimes ne se transforment en incidents sérieux pour le système. Les *near-misses* qui en résultent peuvent fournir un éclairage sur une alternative puissante à la prévention de l'erreur humaine, à savoir la promotion de la récupération de l'erreur par l'opérateur humain. Cette contribution discute diverses approches théoriques pour modéliser la récupération d'erreur et pour les traduire en termes de questions de recherche empirique. Bon nombre de recherches pilotes répondent partiellement à ces questions. Leurs principales conclusions conduisent à penser que la récupération des erreurs est bien plus que le fait de la chance ou de la coïncidence, que ses causes profondes peuvent être identifiées et devraient avoir des implications dans la conception du contexte technique et organisationnel des tâches des opérateurs humains.

MOTS CLÉS : Récupération d'erreur, Fiabilité humaine, Analyse d'incidents, *Near-misses*

COGNITIVE SIMULATION OF A FLIGHT CREW DURING THE APPROACH TO LANDING PHASE

M. Pedrali, M. Fabbri

European Commission
Joint Research Centre
Institute for Systems, Informatics and Safety
21020 Ispra (Va), Italy
mauro.pedrali@jrc.it – marco.fabbri@jrc.it

ABSTRACT

This paper addresses the implementation of a human reliability method for studying the interactions between the crew, the aircraft and the environment within a cockpit, both in nominal and abnormal conditions. The method has been applied prospectively and implemented into a simulation tool for qualitative analyses of the Approach To Landing phase procedure. Errors that can be performed by the crew have been studied, classified by means of a taxonomy of human errors provided by the method, and their consequences assessed. In this paper, the method is presented and its application discussed, focusing on the pilots' erroneous cognitive process.

Keywords: prospective analysis, human errors, flight procedures.

INTRODUCTION

An essential component of mobility is transport by air; one of the primary conditions to transport by air is safety. Improving the level of safety of the Air Transport System (ATS), which is already very high, is one of the most challenging goals in this domain. Safety as perceived by the travelling people is determined by the number of accident. Accident investigations have demonstrated how unwanted occurrences are often the result of mishaps in the human interaction with the physical and socio-technical environment. In order to prevent these events from occurring and contain their adverse consequences, new systematic approaches dedicated to the so-called human "Error Management" are needed (Reason, 1997), based on methods for prospective analyses of human interaction.

A fundamental standpoint to develop such methods, is the availability of a Human Interaction model. This paradigm enables to formalise human behaviours, the system and the environment responses in such a way that hypothetical case studies and realistic working situation can be analysed and formally represented (DoD, 1984). Generally speaking, these interactions always occur in a realistic context which is characterised by the system under control, by the socio-technical context and by the operator(s) in direct contact with the system. While the system interacts with the operator(s) through its interfaces (i.e., display panels, indicators), the socio-technical context affects the working environment influencing human behaviour. Eventually, all these factors may modify the unfolding of the cognitive processes as well as the performance of manual or control actions by, for example, causing inappropriate behaviours or by altering the amount of knowledge accessible to the human in a given circumstance.

This paper discusses an approach based on simulation techniques that can be used for studying potential accidental scenarios occurring in the ATS, where the human element, i.e., the pilot(s), plays an important role. In order to analyse these scenarios, an integrated simulation that implements an interaction model was developed, representing dynamically the relationships between an aeroplane, its crew and the surrounding environment. In particular, the following elements will be expanded as follows:

1. Operators model: Cognitive behaviour of a two-pilot crew, focusing on erroneous actions.
2. Procedures: Set of activities and rules performed by the crew when following a procedure.

These components have been implemented into a computer based simulation and a number of case studies will be shown, addressing the human error issue in the Approach To Landing phase procedure. From these examples, the advantages of using simulation as a support tool will be demonstrated and future development and practical applications of the proposed approach will be considered. Indeed, virtually limitless sequences of events, combining human errors and system malfunctions can be studied through these applications, being only bounded by the adequacy of the simulation.

METHOD FOR PROSPECTIVE ANALYSIS

The selected human reliability method is based on a (1) reference model of cognition and a (2) classification of erroneous actions (Hollnagel, 1998).
The reference model of cognition has been selected among those belonging to the contextual model set. These models consider human cognitive processes as driven both by the stimulus perceived by the operator (in this case the pilot) and by the context (system, procedures, environment) in which the actions take place. In particular, the perceived stimulus represent the starting point of the cognitive process while the context allows the selection of the available cognitive functions and defines the links between them. The combination of the two elements identifies the cognitive process, which in turn, determines the action to be carried out. Although contextual models are, at present, largely accepted by those who study Human Factors in system control, a review of cognitive simulations developed over the last twenty years, showed that contextual models are still rarely used in practical applications [Cacciabue, 1998].

The COntextual COntrol Model, COCOM (Hollnagel, 1993) consists of two components or modules, these are:

1. Competence, describing the ability of the operator in performing a specific set of actions and comprising:
   - The Template Set, principles for the use of the main cognitive functions (Perception, Interpretation, Memory, Planning and Execution) in relation to the context of operations and the external stimuli.
   - The Activity Set, actions an operator, e.g., a pilot, is capable to perform and that are significant in the existing context.
2. Control, indexing the operator competence on the basis of the subjectively available time. This time depends on different factors such as the system dynamics with which the operator interacts, the external environment, the operator's experience and skills, and the availability of external resources necessary for performing the task.

The concept of control is peculiar to this model. In COCOM a cognitive process depends on the control which, in turns, is affected by the context. The control can be represented as a continuum, uni-dimensional, whose limits are characterised respectively by a very low control and a high control. For the sake of simplicity, the model refers to four control modes (listed, respectively, from the lowest to the highest control), namely: "scrambled", "opportunistic", "tactical" and "strategic".

The model is straightforwardly connected with a classification of human errors and with a criterion for its use. The classification makes a clear distinction between the causes of erroneous actions, called genotypes, their effects and manifestations, called phenotypes. Effects and manifestations are respectively the erroneous forms of a cognitive process and the actual external expressions of an erroneous behaviour. The classification is intended to establish a sound structure enabling fully bi-directional analysis, and therefore allowing both retrospective analysis and performance prediction of human interaction (i.e., prospective analysis). While in the first case the aim is to trace back a cognitive process (from the manifestation of an error up to its root causes), in the second case the objective is to speculate about possible manifestations of an erroneous behaviour starting from a set of triggering causes.

The classification scheme relies in COCOM to organise its categories describing, through possible patterns, the relationship between causes and effects of human erroneous actions. According to the classification scheme, the application is recursive rather than strictly sequential (hence it cannot be considered as mere error taxonomy). This feature fits with the underlying reference model, allowing to represent cyclic processes. Finally, a clear stop-rule determines when a prediction or an analysis has come to the end; otherwise, since the criterion is recursive, an analysis or prediction could go on forever.

THE CREW MODEL DESIGN

Through the control module and the mode of control concept, COCOM indicates how to deal with the competence with respect to the context. According to the theory (Hollnagel, 1993), the control is affected by the "outcome of previous actions", the "estimation of subjectively available time" and "sudden changes". However, no specific or explicit indications are provided either on how the independent parameters affect the mode of control, or, conversely, any relationships between the mode of control and the parameters.

Hence, we used the workload to which the crew members are subjected as the independent variable for indexing the mode of control. The workload is considered as a variable measuring the adaptation between the context-of-operation stimuli and the capacity of the operator to deal with it (Kantowitz, 1986). On the basis of this definition, we decided to evaluate the operator workload as the ratio of the following time intervals:

- **Time to Perform** (**Δt_{perfn}**), defined as the period needed by the operator to carry out the n-th action.
- **Time Available** (**Δt_{avl}**), defined as the interval between the moment when the n-th action (T_{sn}) is performed, and the condition for performing the n-th+1 action (Ts_{n+1}).

These intervals are characterised by "Signs" and depend on the dynamics of the system (i.e. the aeroplane) and the context evolution. The workload WL_{n+1} experienced by the operator during the execution of the n-th+1 action is therefore determined according to the following equation, whose entities are illustrated below.

$$WL_{n+1} = \frac{\Delta t_{perfn}}{\Delta t_{avl}} = \frac{\Delta t_{perfn}}{T_{sn+1} - T_{sn}}$$

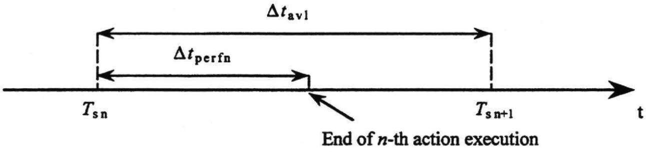


Figure 1 – Entities defining the workload.

The workload values resulting from this equation are as follows:

- Constant, within the time interval occurring between the Sign n-th and the Sign n-th+1.
- Dimensionless.
- Always included in the interval [0, 1] (assuming that each pilot performs one action at a time).

Afterwards, we define the relationships between the workload and the control mode of the operator. They are based on an inverse proportion between the workload WL_{n+1} and the control mode related to action n-th+1. Hence, high workload values correspond to low control modes and vice versa. As COCOM foresees four distinct control modes, the continuous interval [0, 1] has been divided as follows:

$$\begin{cases} WL_{n+1}=1 \text{ and } NNC=1 & \Rightarrow \text{Control mode "Scrambled"} \\ a_1 \leq WL_{n+1} \leq 1 & \Rightarrow \text{Control mode "Opportunistic"} \\ a_2 \leq WL_{n+1} < a_1 & \Rightarrow \text{Control mode "Tactical"} \\ 0 \leq WL_{n+1} < a_2 & \Rightarrow \text{Control mode "Strategic"} \end{cases}$$

Values 'a_1' and 'a_2' define the workload ranges corresponding to each control mode; the numerical value of 'a1' and 'a_2' depend on the context of operation and must be selected within the interval (0, 1). In above expressions the variable NNC (Non Normal Conditions) has been introduced to emphasise the extreme character of the "Scrambled" control mode. NNC has been conceived as a binary variable that is set to 1 each time the system and/or the environment do not behave as the operator expects.

THE CREW MODEL IMPLEMENTATION

The crew model has been implemented into a simulation to carry out prospective analyses of the Approach To Landing procedure (Fabbri, 1998); the simulation was developed through the following design phases:

1. Context Description. The Approach To Landing procedure towards a specific airport was considered and the aircraft was assumed to operate in nominal conditions.
2. Task Analysis. The set/list of activities performed by the crew to fulfil the ATL procedure was examined.
3. Errors Identification. Each activity was decomposed into its related actions and likely errors investigated.
4. Performance Prediction. The relationships between the manifestation of erroneous actions (phenotypes) and their triggering causes (genotypes) were hypothesised as follows:
   - θ Definition of the likely phenotypes.
   - θ Identification of the causal pathways (through the classification) leading to the phenotypes.
   - θ Indexation of the current crew member control mode.

In the following paragraphs we will detail the most characterising phases of the implementation process, namely: the Errors Identification and the Performance Prediction.

Errors Identification in the ATL Procedure

The activities that each crew member has to perform depend on the navigation procedure (in this case the ATL) and the type of aircraft (in this case a Boeing 747-200). Examples are flaps and landing gear extension, altimeter calibration, each activity is characterised respectively by:

- θ Signs
- θ Actions

While Signs are those conditions the system, i.e., aircraft, has to reach in order to make the pilot perform a given action. Actions are all the operations that have to be accomplished by the each crew member involved. In order to investigate the possible erroneous actions related to the activities, we decomposed them using the event tree technique, typical of the THERP formalism (Swain & Guttmann, 1983).
Briefly, each *Activity* was associated to the corresponding *Actor*, i.e., Pilot Flying (PF) and Pilot Not Flying (PNF), and subdivided in a number of *Acts*. For each activity (13 altogether), a graphic representation was provided enabling to identify the likely errors.

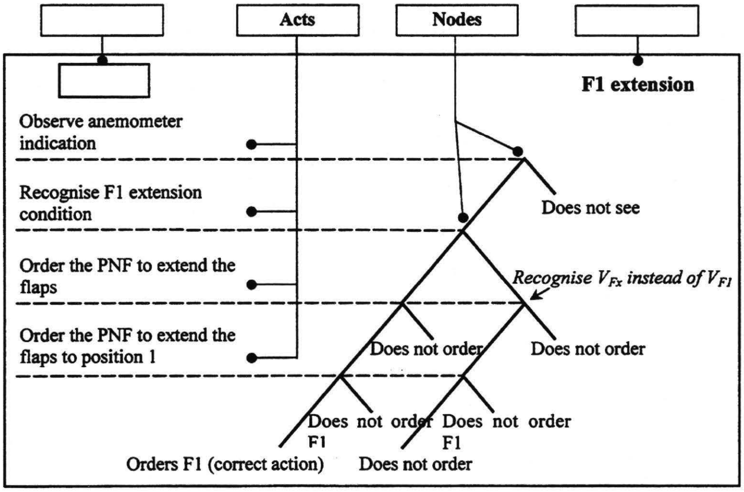


Figure 2 – Study of the possible errors related to the flaps extension.

Performance Prediction

This is the characterising feature of the prospective analysis, whose peculiarity is dynamics. Although the likely phenotypes and causal pathways can be determined a priori for a given control mode index, this index actually depends on the context. The context constantly changes and usually in an unpredictable way because of the system dynamics. This feature was completely implemented in the simulation tool.

A specific algorithm determines the possible patterns that the error classification can assume according to (1) the flight crew workload and the cognitive functions involved, (2) the possible phenotypes, and (3) the description of the context in which the pilots operate. These parameters identify the possible cause-effect links within the classification scheme leading to a phenotype. The crew member erroneous action is determined by matching this phenotype with the context-related error, which is specific to the activity performed at that moment. This index depends on the workload as mentioned previously. It is interesting to note that the mode of control 'Tactical' does not foresee the use of the Planning function.

APPLICATIONS

The model was implemented into a computer code and it was developed following the object-oriented approach, in particular the programming language C++ was used (Stroustrup, 1986). The object-oriented methodology was selected mainly because it helps structuring data "as they are" in the real application domain, and it allows to organise the simulation code as a collection of stand-alone modules. To this regard, the 'cognitive' simulation was embedded in a more general simulation frame in which the aeroplane model reproduces the system performance and provides the crew through an interface model (the cockpit instruments) with the information needed to perform the selected procedure (Barbieri & Vallana, 1998).

A number of scenarios were investigated by running the simulation code; in this paper three case studies are discussed. While the first focuses on the differences between COCOM and a procedural cognitive model, highlighting COCOM advantages, the others investigate crew performance matters, focusing on consequences of pilots' errors on the ATL procedure execution. In all these cases, hypotheses were made on aircraft conditions, assuming that it properly operates without any system failure. In the following each case and the related results will be discussed.

The objective of first case was to evaluate the workload model in nominal conditions, and therefore neither crew errors nor particular environment conditions were simulated during the entire ATL phase. Results show the crew workload variation vs. time. The workload between tasks n-th and n-th+1 is constant, therefore each workload curve step corresponds to a task execution. Each crew member undergoes two workload peaks having different duration in time; these peaks are generated by the task sequence occurring within a short time interval and depending on the aeroplane dynamics (Figure 3).

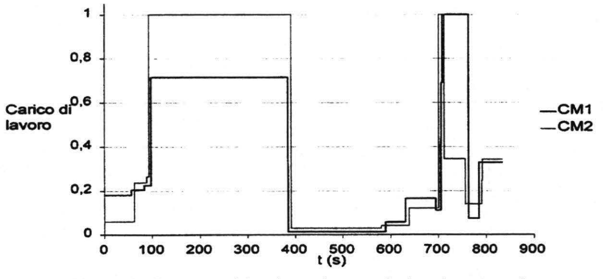


Figure 3. Crew workload vs. time variation (no errors).

In the second and third case the error classification has been applied. In particular the crew is supposed to err when performing the Altimeter Calibration and the Approach Check List (altimeter check) It was decided to focus on these tasks because they are critical while performing an ATL procedure. As a matter of fact the correct altimeter setting allow the crew to maintain (a) an appropriate vertical separation from the ground (terrain, obstacles) and (b) a correct descent profile. The Approach Check List item "Altimeter check" is deemed to verify the altimeter indication and if properly executed allows to recover from an eventual calibration error. Simulation results show that the Pilot Flying (PF) calibrates the altimeter on a value (1008 hPa) different from the reference sea level pressure set as initial condition value (1003 hPa) and that the Pilot Not Flying (PNF) does not perform the altimeter calibration check included in the approach check list. The simulation stops when the PF performs the radio-altimeter check. Indeed the navigation procedure requires the crew to interrupt the approach sequence when the absolute difference between the altimeter and the radio-altimeter readings are greater than a determinate value (23 m) the crew workload variation vs. time is shown in Figure 4.

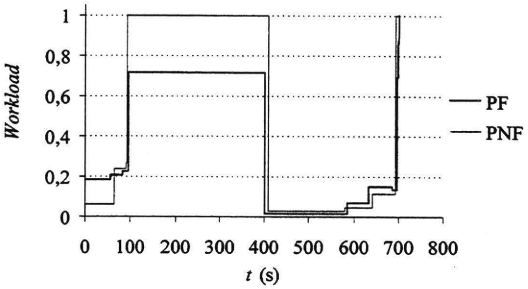


Figure 4. Crew workload vs. time variation (with errors).

Results reveal the same workload values experienced by the crew in the nominal performance, owing to the fact that, when doing an erroneous action, both the PF and the PNF behave as they were acting correctly.

CONCLUSION

Prospective analyses of the human interaction in the air transport system have become essential. This kind of approaches allow both to predict hypothetical hazardous events and to introduce safety measures in order to prevent these events from occurring in real flight operations. Two main advantages came out from this study. The first one is that variety of cases that can be evaluated and tested at a very low price and in relatively short time. The second derives from the use of a contextual and control model combined with a classification of erroneous actions, highlighting the potential from the application of such a method to domains where the role of operators (e.g., pilots) is critical with regards to operations safety.

REFERENCES

Barbieri, A., & Vallana, S. (1998). *Il wind-shear nella sicurezza del volo: Sviluppo di una simulazione integrata velivolo-pilota-ambiente*. Graduation Thesis in Aeronautical Engineering (in Italian), Politecnico di Milano, Italy.

Cacciabue, P.C. (1998). *Modelling and Simulation of Human Behaviour in System Control*. London: Springer-Verlag.

Fabbri, M. (1998). *Metodi di analisi prospettica per lo studio di errori umani: Simulazione cognitiva dell'equipaggio di un velivolo durante la fase di avvicinamento*. Graduation Thesis in Aeronautical Engineering (in Italian), Politecnico di Milano, Italy.

Hollnagel, E. (1993). *Human Reliability Analysis – Context and Control*. London: Academic Press.

Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method*. London: Elsevier.

Kantowitz, B.H. (1986). Mental workload. In P.A. Hancock (Ed.), Human factors psychology. Amsterdam: North Holland.

Stroustroup, B. (1986). *The C++ Programming Language*. Reading, MA: Addison-Wesley.

Swain, A.D., & Guttmann, H.E. (1983). Handbook on Human Reliability Analysis with Emphasis on Nuclear Power Plant Application (Draft Report CR-1278). Washington D.C.: NUREG.

RÉSUMÉ

SIMULATION COGNITIVE D'UN ÉQUIPAGE D'AVION PENDANT LA PHASE D'APPROCHE À L'ATTERRISSAGE

Cette communication traite de la mise en œuvre d'une méthode d'analyse de la fiabilité humaine pour l'étude des interactions dans le cockpit entre l'équipage, l'avion et l'environnement, à la fois dans des conditions normales et anormales. MOTS CLÉS : Analyse prospective, Erreurs humaines, Procédures de vol

COGNITIVE MODELING OF SHIP NAVIGATION AND ITS APPLICATION TO RISK ANALYSIS

Kenji Itoh
Tokyo Institute of Technology
e-mail: ken@ie.me.titech.ac.jp

Toshinori Yamaguchi
Tokyo Institute of Technology
e-mail: yamaguti@ie.me.titech.ac.jp

John Paulin Hansen
Space Invaders
e-mail: paulin@invaders.dk

Finn R. Nielsen
Risø National Laboratory,
e-mail: finn.nielsen@risoe.dk

ABSTRACT

The present study examines the possibility of a simulation approach with cognitive modeling to risk analysis in ship navigation. We constructed a *cognitive model* of a ship navigator for a simple course tracking task based on the cognitive task analysis in which synchronized data from eye-movement recordings and verbal protocols were analyzed during a series of experimental sessions using a maritime simulator. The cognitive model was confirmed by the simulation results of ship motions and navigator's cognitive behaviour using the identical scenarios to the experimental navigation sessions. Then, we applied the cognitive model to risk analysis by performing a series of simulation runs under various navigation scenarios changing individual factors and manoeuvring conditions. As one of the results of risk analysis, it was found that a navigator's poor competence such as identification and perception errors is a critical risk under strong current conditions.

Keywords: Ship navigation, cognitive modelling, cognitive simulation, risk analysis

INTRODUCTION

The human cognitive processes play a crucial role for the safety of maritime operations. The crew's competence is of considerable importance to manoeuvring, especially when environmental conditions confine the safety margins (Schuffel, 1986). A *simulation approach* has been utilized for its advantage of cost and time savings to investigate risks under specific manoeuvring conditions (e.g., Papenhuijzen, 1988) like in other fields (e.g.. Cacciabue et al., 1990). In this approach, modelling of human behaviour and performance is of primary interests. However, most of the existing ship handling models are based on control theoretical approaches rather than on the cognitive performance of human navigators. This is due to the difficulty of analyzing and modeling navigator's cognitive processes which are characterized skill- and rule-based performance (Rasmussen, 1986).

The present study presents a simulation approach with cognitive modelling to risk analysis in the ship navigation. A ship navigator's *cognitive model* was built for a simple course tracking task based on task analysis using a maritime simulator (Itoh, et al., 1998). We examine the cognitive model's descriptive and predicitve capabilities by comparing simulated ship motions and the navigator's behaviours with those in the human navigators using the identical scenarios to the experimental navigation sessions. After confirmation of the model. we apply the simulation approach to risk analysis by performing a series of simulation runs under various navigation scenarios changing individual factors and manoeuvring conditions.

NAVIGATOR'S COGNITIVE MODEL

Modelling of Course Correction Process

The cognitive model represents the dynamic interaction between the navigator's behaviour and the states in manoeuvre during performance of course tracking task. His actions are controlled by the stochastic transition network of eye-movements which was obtained by the cognitive task analysis (Itoh et al., 1998). The track plots and the navigator's mental states during task performance can be estimated by simulation of the model under a specific scenario. All the modules were programmed in standard C code, and implemented on a Macintosh PC.

The navigator model controls the gaze-transition during task performance. The eye-gaze patterns, i.e., transition probabilities between any two information sources and gaze time at each source, were obtained by the task analysis both for state monitoring and for course correction operation (Itoh et al., 1998). Gazed information at each point and its gaze time are determined by the random digit according to the stochastic distributions in a simulation run. The indication of the gazed information source is obtained from the ship motion module and is saved in STM(short term memory). The present position and heading are identified when the model's "eye" gazes the outside scene or the radar picture. The accuracy of obtained information is biased according to the assumed competence of the navigator, which is installed as parameters representing individual human factors. A model of poor navigator has larger error factors for identification of the heading and vessel position.

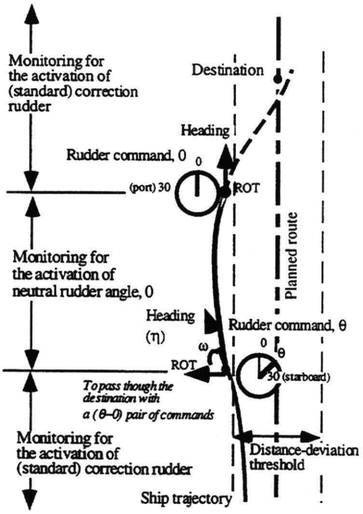


Figure 1 Outline of cognitive process in course tracking

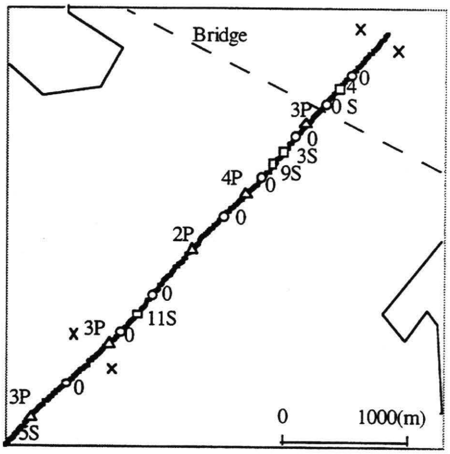


Figure 2 Track plot navigated by the cognitive model

As the outline of the navigator's process is depicted in Figure 1, the model repeatedly checks the deviation in course and position from the planned states during the state monitoring. In the daylight condition, for example, such a *deviation check* for the course correction is performed every time the outside scene is attended to by the model's "eye". When the present course deviation, the present or future distance deviation is identified to be bigger than its threshold, a course correction operation is activated. Each threshold value varies depending on individual preference and competence among navigators, the environmental load, and the manoeuvrability of a particular ship. These values are treated as variables in the model. Considering the manoeuvre in the narrow fairway, the distance and course thresholds in the simulation runs are set for a skilled navigator's model to 50m and 2.0 deg. for each side of vessel, respectively.

When an operation is activated, the navigator model tries to generate *a pair of* rudder commands, i.e., a non-zero rudder and a subsequent neutral rudder settings, so that the ship can sail to the destination for course correction smoothly with desirable heading and ROT (rate-of-turn). The navigator model determines a rudder setting for the first non-zero command with reference to state variables such as the present heading and ROT in the following sequence. First, a destination for the course correction is generated as the intersection of the pilot line and the planned route. Then, the destination is modified, taking into account the effect of the wind and/or current. The present model treats only the current, and the navigator's identification error of current strength is also modelled as a relative error factor. The modified destination is obtained by offsetting the distance vector, to which the ship is carried away by the identified strength of current, from the original destination point. As the navigator is assumed to have a mental model on simplistic ship dynamics, the decision process of rudder setting for the modified destination is implemented as a quick reference of the decision table with the necessary state variables.

After the non-zero rudder is given, the navigator model performs state monitoring for the timing decision on returning the helm to the neutral position, employing a *generate and test* procedure. The future track is *generated* by applying a *look-ahead* function to the present manoeuvring state with the helm operation returning to the neutral position. The predicted future position is *tested* whether the expected course is on the destination within an allowance limit. If the test is satisfied, then the model accepts the present timing to issue a zero-degree command. Otherwise, the *next* point is generated when the model's "eye" fixates the outside scene. This process is repeated until a satisfied timing is obtained, every time the outside scene is gazed in the daylight manoeuvring. When the generated future track crosses the planned route at near side to the destination, the timing is decided too late to ease the helm. Then, the model is going back to the normal state monitoring process. In this case, the helm angle with non-zero degrees is given at the next operation, instead of returning it to the neutral position.

Modelling of Human Errors

The navigator model is equipped with a simple *human-error process*: misidentification of the present position and heading in monitoring the outside scene and the radar screen, misidentification of current strength, and command generation error. Each of these errors is characterized by a probabilistic distribution dependent upon the navigator's competence. The cognitive model has two parameters to represent an perception error of heading or position: mean bias from the correct value, and its standard deviation in the absolute error. Only a single parameter, i.e., relative error to a desirable rudder, is applied to modelling inaccuracy of a generated command. For example, two random numbers are selected to determine identification errors of the present position and heading, anytime the model performs monitoring of outside scene. The model calculates navigator's perception errors of heading and position following their prescribed stochastic distributions, and offsets them from correct values to determine perceived heading and position.

MODEL VALIDATION

Task Network Simulation

The task network simulation of the navigator model was executed under the identical conditions to the daylight and the night-fog course tracking sessions to examine the model's descriptive and predictable capabilities toward navigator's cognitive processes and ship motions. The navigator's action sequences and course plots simulated by the model are compared with the observed performance by the human navigator in the experimental sessions (Itoh et al., 1998). A simulation run enables to reproduce and display on a PC screen not only a track plot and ship motion but also the navigator's action sequences and transition network of attention during state monitoring.

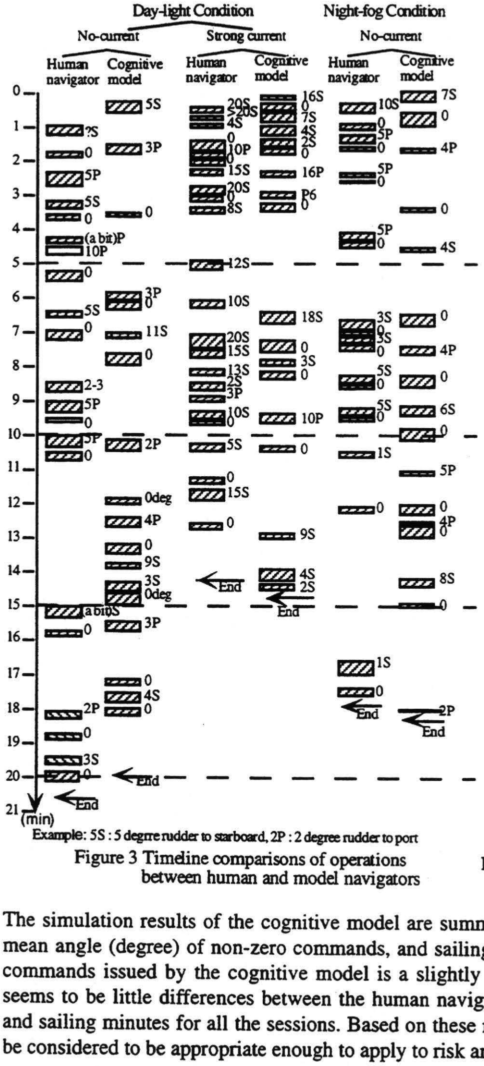
Simulation Results

In both the daylight and the night-fog conditions, the ship was navigated by the cognitive model safely following the planned route. Figure 2 depicts an example track plot of manoeuvring in the no-current daylight session. This track plot is found to be similar to the one obtained by the human navigator in the simulator session (Itoh et al, 1998) not only in the number of commands and their rudder settings but also in the command sequence. In addition to the track plot, the time lines of operations are presented for all the three sessions in Figure 3. This figure includes the operation sequences produced both by the human navigator and by the cognitive model to compare easily their command generation patterns. Under the no-current conditions in the daylight and night-fog sailing, it is found that the model employed a "return to neutral rudder" strategy like the human navigator. After a non-zero rudder command for course change, the model and human navigators returned the helm to the neutral position in most cases. In contrast, this strategy was not employed by both the human navigator and the model to manoeuvre in the strong current condition. Instead, both navigators generated a series of starboard setting with a few zero rudders due to strong current.

Table 1. Performance comparisons of experimental sessions between human navigator and cognitive model

	Daylight & no-current			Daylight & strong current			Night-fog & no-current		
	Human op.	Simulated op.		Human op.	Simulated op.		Human op.	Simulated op.	
		mean	SD		mean	SD		mean	SD
No. of operation	21	17.0	2.3	23	18.2	3.4	20	16.5	2.4
Rudder angle(deg.)	5.0	4.9	0.6	11.9	8.4	2.0	4.2	4.9	1.0
Sailing min.	1230.0	1161.3	2.5	840.0	920.5	6.5	1080.0	1115.0	23.4

(Averaged value over 10 simulation runs with diffrent random digit seeds for each condition)



Example: 5S : 5 degree rudder to starboard, 2P : 2 degree rudder to port

Figure 3 Timeline comparisons of operations between human and model navigators

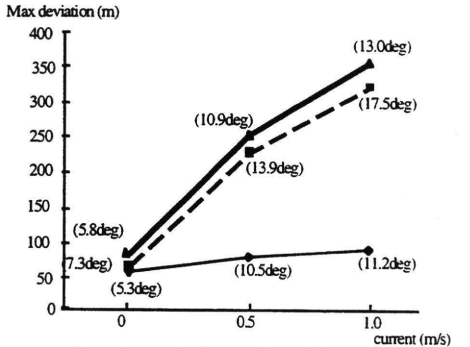


Figure in bracket(): Mean rudder angle in operations

(a) Maxmum deviation and rudder angle

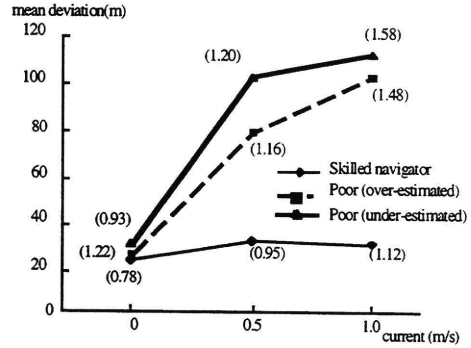


Figure in bracket(): Mean operation frequency operation (No. ope./min.)

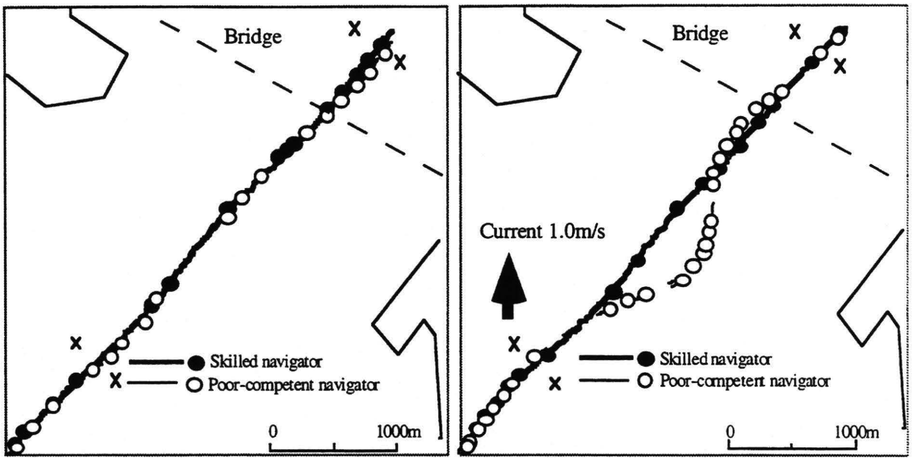
(b) Mean deviation and operation frequency

Figure 4 Simulated performance of skilled and poor-competent navigators under different current conditions

The simulation results of the cognitive model are summarised in terms of the number of performed operations, mean angle (degree) of non-zero commands, and sailing time in Table 1. From this table, the number of rudder commands issued by the cognitive model is a slightly less than that by the human navigator. However, there seems to be little differences between the human navigator and the model in the mean angles of rudder setting and sailing minutes for all the sessions. Based on these results, the cognitive model constructed in this study can be considered to be appropriate enough to apply to risk analysis.

APPLICATION TO RISK ANALYSIS

In order to apply the cognitive modelling approach to risk analysis, a series of simulation runs are conducted using various navigation scenarios, e.g. changing environmental conditions and navigator's individual factors. Example risks to be examined here are strong current and a navigator's poor competence. The levels of current during manoeuvre were varied between zero and 1.0m/s in the same direction. The navigator's competence was modeled as error distributions in identifying heading, position and current as well as an error distribution of command generation, as mentioned previously. These error parameters for a skilled navigator were set to zero biases with 1.0 deg. and 10m of standard deviations (SD) for identification of heading and position, respectively. He had a 20% reading error of current less than the actual strength, and no error generation of command setting. We exemplified two poor-competent navigators whose parameter settings were plus and minus 2.0 deg. and 15m biases with 1.5 deg. and 10m SD's for identification errors, and plus and minus 40% reading error of current and 30% relative error of command generation in SD.



(a) No-current condition (b) Strong current condition

Figure 5 Simulated track plots by skilled and poor-competent navigators

Maximum and mean deviations from the planned route as well as mean rudder setting and operation frequency are shown for each simulation condition, i.e., current conditions and navigator's competence levels, in Figure 4. As can be seen in this figure, navigation by the skilled navigator's model seems to be stable in terms of deviation from the ideal route and operation frequency under any current condition. In contrast, when the poor-competent navigator's models navigate a ship, sailing tracks are affected largely by the current. The distance deviation from the ideal route becomes bigger with the strength of current. For example, the maximum deviation is estimated approximately 350m under 1.om/s strong current. Such a large deviation is very dangerous in a narrow fairway like Oersund channel taken up in this study. We performed additional simulation runs varying the error factor of current identification with keeping poor-competent levels for the other error factors. From the simulation results, the identification error of current is suggested to be a major source of the big deviation from the ideal track. It is also found that poor competence of a navigator causes frequent manipulations of steering wheel to follow the route. To easily understand effect of navigator's competence to the sailing track, ship trajectories simulated by a skilled and poor-competent navigator's models are shown in Figure 5. These simulation results suggest that poor competence is a critical risk for loosing control of ship especially in difficult manoeuvring conditions.

CONCLUSION

This study constructed a cognitive model of a ship navigator for course tracking operations based on the task analysis of some manoeuvring sessions using a ship simulator. We applied this model to risk analysis through a series of the man-in-the-loop simulation with varying manoeuvring scenarios and navigator's individual factors. Based on the simulation results, it is found that the navigator's poor competence is a critical risk for manoeuvre in strong current conditions.

We are now performing over-night batch simulations with various manoeuvring scenarios, e.g., sailing course and fairway condition, and navigator's individual factors such as look-ahead time window and its accuracy, threshold of course and distance deviations, applying the cognitive modelling approach proposed in this study. To facilitate such automated risk analysis, we plan to integrate analysis features of simulated data log of ship motions and human behaviour with the cognitive simulation program in a future project. For example, we are designing several analysis modules: calculation of mean square error of sailing track from the planned route, time line analysis of navigator's action and transition network of the navigator's attention. In addition, it is also of great importance to integrate the evaluation of the navigator's mental workload during task performance with the cognitive simulation. We believe it is possible to estimate the mental workload based on the STM states which can be speculated by a cognitive model (Itoh, 1998).

Acknowledgment

This work was in part supported by by Grant-in-Aid for Monbusho International Scientific Research Program (No. 80159871), the Japanese Ministry of Education, Science and Culture. Leif Løvborg and Steen Weber. Risø National Laboratory and Jens Rasmussen, HEURECON all provided useful ideas to the project. Special thanks to Jakob Buus Petersen, The Danish Maritime Institute, for his participation and suggestions.

REFERENCES

Cacciabue, P. C., Decortis, F., Drozdowicz, B., & Kjaer-Hansen, J. (1990). A framework for cognitive simulation of human decision making and behaviour in complex work environments. In H.B. Andersen. S.A. Pedersen, C. Cacciabue, & J. Reason (Eds.), *Cognitive Processes and Resources, Proceedings of the second MOHAWC Workshop* (Vol. 1, pp.25-58), Manchester, UK, December.

Itoh, K. (1998). Cognitive modelling scheme for task analysis and its application to computer editing. *International Journal of Cognitive Ergonomics*, 2(3), 269-295.

Itoh, K., Hansen, J. P., & Nielsen, F. R. (1998). Cognitive modelling of a ship navigator based on protocol and eye-movement analysis. *Le Travail Humain*, 61(2), 99-127.

Papenhuijzen, R. (1988). On the modelling of the behaviour of a navigator: background and proposal for a control theoretic approach. In J. Patrick, & K .D. Duncan (Eds.), *Training human decision making and control* (pp. 189-203). Amsterdam: Elsevier.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York: North-Holland.

Schuffel, H. (1986). *Human control of ships in tracking tasks*. Doctoral dissertation, University of Tillburg. The Netherlands.

RÉSUMÉ

LA SIMULATION COGNITIVE DE LA NAVIGATION MARITIME ET SON APPLICATION À L'ANALYSE DES RISQUES

La présente étude examine la faisabilité d'une approche de l'analyse des risques dans la navigation maritime par la simulation cognitive. Nous avons construit un modèle cognitif de l'officier de quart dans une tâche de suivi de route, à partir d'une analyse de données synchronisées recueillies au cours d'une série de sessions expérimentales sur un simulateur maritime : les mouvements oculaires et les verbalisations. Le modèle cognitif a été validé par les résultats de la simulation sur les mouvements du navire et par la confrontation à l'activité cognitive d'un officier sur les mêmes scénarios. Ensuite, nous avons appliqué le modèle cognitif à l'analyse des risques en réalisant une série de simulations sur des scénarios variés, en changeant des facteurs individuels et des conditions de manœuvre. Parmi les résultats de l'analyse des risques, nous avons pu montrer que la faible compétence de l'officier, se traduisant notamment par des erreurs d'identification et de perception, est un risque critique quand les conditions de manœuvre sont difficiles.

MOTS CLÉS : Navigation maritime, Modélisation cognitive, Simulation cognitive, Analyse des risques

ON MODELS OF INCIDENTS AND ACCIDENTS

Ola Svenson

Department of Psychology Stockholm University, S 106 91 Stockholm,
Sweden e-mail: Ola.Svenson@psychology.su.se

ABSTRACT

An accident can be explained in different ways depending on the accident analysis model that is used. Some fundamental accident models are presented here. They can apply in isolation and in combinations with each other. These generic models are associated different recommendations for increasing the safety of a system. If an analyst or an authority subscribing to the analysis has a clear understanding of the fundamental models and perspectives underlying an accident analysis, his or her evaluation of the conclusions will be more insightful and less biased than if she or he did not have these insights. In addition, explaining an accident is different from predicting an accident and human cognition is very poor at making this distinction. Therefore, the risk of drawing improper causal conclusions from accident analyses can be derived both from fundamental accident models and from the human shortcoming not to make a proper distinction between the explanation of an accident and the prediction of an accident (needed for successful actions to avoid future accidents).

Keyords: accident, incident, accident analysis, risk analysis, human factors.

INTRODUCTION

Lay people and experts use a number of intuitive and or analytical models of how accidents are caused. Models from everyday life have a tendency to be more basic and intuitive than the more sophisticated models that are used in professional incident and accident analyses. Still, everyday life simple intuitive models are also part of professional incident and accident analysis models. The present contribution will illustrate these simple intuitive models that are used in more complex models of incidents and accidents. The meta perspective chosen here is that an accidents is the result of malfunctions of some system(s) and/or system interactions. Given this, one main problem facing an accident analyst is that incidents and accidents in complex systems are complex and that accident analyses have to rely on simplified models of this complex reality.

Individuals learn to modify their behavior through the errors they make if proper feedback from the errors is provided and if this feedback is interpreted correctly. In the same way, people and organizations can learn from accident analyses. However, it is important to point out that it is not always that the most effective ways of coping with difficulties are learned from such analyses. What is learned from an incident or an accident depends on the model used for analyzing the accident. As mentioned above models reflect simplifications from one or the other perspective. This is linked to the fact that each model is more or less "wired" to favor one or the other set of causal explanations for an accident.

In addition to different models, there are different perspectives in which accidents can be regarded, such as, legal, technical, human factors, economic, organizational perspectives. The perspectives are related to different generic models explaining why incidents and accidents occur and to different preferred means to avoid such events in the future. To illustrate, an organizational perspective accident analysis leads to an emphasis on organizational explanations and organizational changes to increase safety. A legal perspective leads to an analysis with the purpose of finding one or more persons who can be convicted.

SOME GENERIC MODELS OF ACCIDENTS

Concepts that can be used in describing accident models are elements, relations, process, systems, time and space. Elements may represent objects, subsystems and information and there are causal and correlational relations. The following will illustrate the use of these concepts.

Single Root-Cause Models

The simplest of all incident and accident models is the Single Root-Cause Model. In this model it is sufficient to find one factor as the cause of the incident or accident. The model provides a framework for information that is very easy to remember and which, if it is convincing, gives a cognitively simple understanding of the world. From a legal perspective this is an attractive model because it focuses on only one cause, a fact that can make it easier to find one or a few persons responsible for the cause. Most people, when asked about it may not accept the single root-cause model, and yet they may apply the model themselves in many contexts.

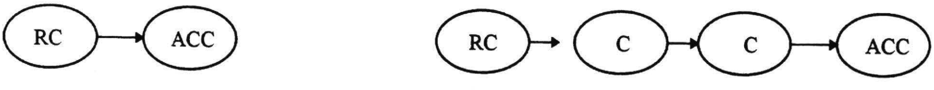


Figure 1 Single Root-Cause Model with one antecedent causing the accident (left) and Extended Root-Cause Model in which the immediate cause of the accident is derived to an earlier root-cause (right). RC = root cause, C = cause, ACC = accident.

Extended Root-Cause Model

It the single root-cause model is extended in time, the extended root cause model emerges. Thus, a more sophisticated, expanded version of the single root-cause model is the model extended backwards in time with a root-cause behind the cause of an incident or accident. In this way a chain is formed with a root-cause linked to the accident through one or more intermediate causes.

Multi-Cause Joint Effects Model

If the single-root cause model is extended in width, allowing a number of factors simultaneously causing an accident instead of just one, then the Multi-Cause Joint Effects Model emerges. In this model an incident or accident is viewed as the result of different single root-cause factors simultaneously contributing to an incident or accident.

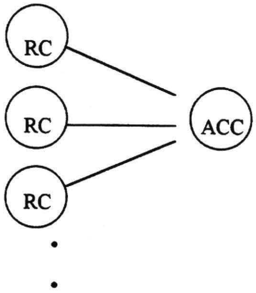


Figure 2 Multi-Cause Joint Effects Model in which a number of factors are related to the accident

The correlational interactions (e.g., facilitation) between the causing variables are of interest in Multi-Cause Joint Effect Model accident analyses. To exemplify, old equipment in conjunction with poor weather may enhance the risks of an accident. The prototype analytical tool for analyses according to this model is multiple regression analysis (used, e.g., in traffic safety research). However, regression analysis can only be used when sufficient numbers of accidents or incidents occur.

Extended Multi-Cause Joint Effect Model

When the width of the Multi-Cause Joint Effect Model is combined with the depth in past time of the Extended

Root-Cause Model the Extended Multi-Cause Joint Model emerges. The causes are traced back in time to a set of root-causes (Figure 3).

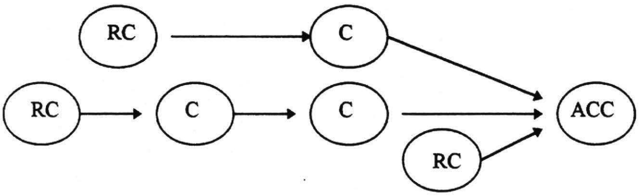


Figure 3 Extended Multi-Cause Joint Effect Model

Causal Tree Model

If causes are allowed to combine not only in a correlational way, as in the Extended Multi-Cause Joint Effect Model, but also in a tree structure, this provides the foundation for a Causal Tree Model.

The typical *Causal Tree Model* links a top event (an incident or accident) to a logical tree of which the branches indicate how the accident was caused by a series of events and nodes (Rasmussen, Duncan & Leplat, 1987). The tree consists of a number of single cause models linked towards the top event through logical gates. Basically, the structure assumes that the different branches leading to the top event are independent if not depicted otherwise.

A *Fault tree* is an example of a logical tree model in which only errors and faults are modeled. In such a model, an accident is caused by a number of faults describing a path with joining branches through the fault tree. The factors in causing an incident or accident are modeled only through "and" and "or" logical gates in the tree.

Event trees start with some event and unfold the possible events that may follow, for example, an accident. The logical structure is identical to that of a fault tree, with the exception that the tree starts with and diverges from the incident/accident instead of converging towards the incident/accident as in a fault tree. An incident or accident analysis according to this model also identifies different possible consequences of an incident/accident even if was stopped in time.

Reason´s Unsafe Acts Model of Accident Causation

Reason (1991) models accidents as the result of a set of failing protective system that usually protect the system from unsafe acts. His "Swiss cheese" model originates in a human factors perspective but can be generalized to a general systems level. Several more sophisticated accident causation models share this kind of generic accident model with a number of protections in depth all of which have failed during a certain accident evolution. This links to the next group of models.

Barrier Function Failure Models

As just noted, most complex systems are designed so that they are protected with defenses in depth against different threats to normal system functioning. To exemplify, a pilot in an airplane is assisted by emergency barrier functions, such as, automatic controls and warnings when some indicator deviates from the normal. The systems performing these *barrier functions* (e.g., adjusting a course, transmitting a warning) are called *barrier function systems.* In barrier function failure models an accident is viewed as a series of errors, where the barrier functions needed all fail or do not exist, leaving the path to an accident free.

Event barrier failure models give full accounts of all the most important events (failures and successes) leading to and contributing to an incident or accident. An example of this kind of model is the Human Performance Evaluation System, HPES (Impo, 1987). This and other models do not separate human organizational and technical errors and barrier function systems. Such a separation is part of the Accident Evolution and Barrier Function (AEB) model (Svenson, 1991).

Fault barrier failure models are focused on faults, deviations and errors that preceded an incident or accident. Events that are not errors are not modeled in models of this kind. In this respect, fault barrier function models parallel the fault tree models, also focusing on errors and failures. An example of a fault barrier failure model is the Accident Evolution and Barrier Function Model, AEB (Svenson, 1991). Figure 4 gives an example of that model depicted graphically.

General description of the accident

A car is driven northwards on a highway. The roadway is wet from a strong shower of rain and there are pools of water on it. The driver remains in the left lane after passing another car. When driving through a right bending curve, he intends to bring the vehicle back in to the right lane. During the maneuver the car skidded and the driver lost control of the car. The car spins round and goes backwards off the road and ends up on the roof. The accident occurred in daylight.

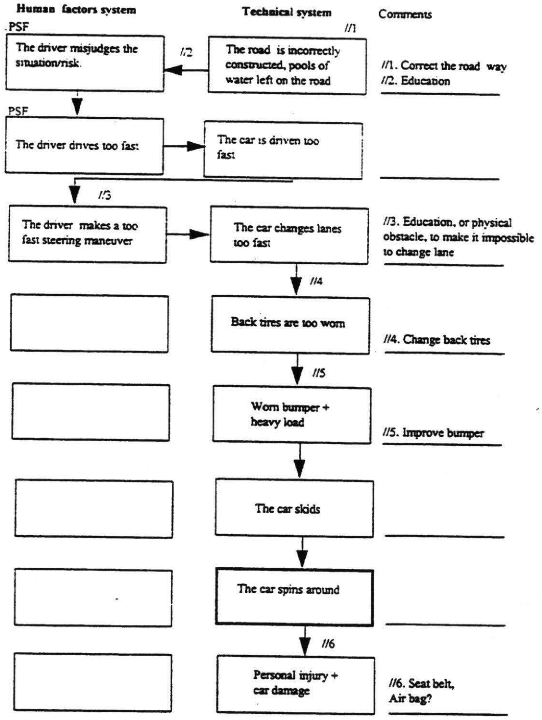


Figure 4 Barrier failure model illustrated by the Accident Evolution and Barrier Function (AEB) Model

How to prevent accidents from the perspectives of the different models

It is quite clear that the above models are prototypes representing oversimplifications and should be regarded only as components in more elaborate analyses. However, it is argued that many recommendations following accident analyses rely on simple models of the kind described above, often embedded in more complex models. There are both theoretical and practical reasons for this. From a theoretical perspective, simple cause effect relationships are part of the models and provide the information needed to influence one ore more causes for improved systems safety. From a practical perspective, recommendations that rely on simple causal models correspond to our intuitive ways of understanding the world and therefore simple model recommendations seem trustworthy and worthy of implementation. When some actions are favored by a model this means that others are less likely to be recommended. Therefore, different models can have different blind spots that may prevent some kinds of actions. The problem of how far back an analysis should go to find root causes is a problem for most models has been highlighted by Hollnagel (1998) but also treated by Rasmussen and Svenson (cf., Svenson, 1991).

In the *Extended Root-Cause Model* the causal chain of errors is traced back behind the immediately causing event. To exemplify, a pilot selecting a too low altitude track before landing and therefore crashing into a hill represents a human error. In the *Single-Root-Cause* approach an obvious action to avoid further accidents is to train pilots to be more attentive when they read the instruments or to fully automatize the pre-landing trajectory. In the extended root-cause perspective, the analysis goes further behind the pilot error searching for reasons why he made the error. This search may result in actions, such as, pilots not allowed to work too long hours and ergonomic change of the altitude meter indicators.

When the *Joint Multi-Cause Factors* model has been selected as a basis for action, regression analysis is an efficient quantitative method for determining the importance of the different factors in the model. By way of example, analyzing a number of factors related to traffic accidents, gives speed of vehicle is a very important role. This has been used to motivate campaigns for reducing the speeds on motor ways. Thus, the quantitative version of the model provides rationales for elimination or change of factors that are seen to contribute significantly to incidents and accidents. It does not explain why but describes the influence of the different factors. Why a factor influences an accident, is a question that has to be addressed using other models.

As presented by Reason, the *Unsafe Acts Model* is more geared towards human and organizational factors than the other models. However, from a theoretical point of view it is equally applicable to technical systems and engineering and can be integrated in *Fault and Event Tree Model* perspectives. When an accident is modeled in a *causal fault tree* or as a path through a *fault tree* this implies that a number of errors have occurred both simultaneously and sequentially. If only one of the error conditions wase not fulfilled, in the path of faults through the tree, the incident or accident would not have happened. The path of faults can be traced backwards to an event called the initiating event. When the path of faults has been triggered, other dormant or latent unknown errors are revealed while the incident/accident develops. The obvious actions following a fault tree model analysis of an incident are to eliminate the initiating event, stop the path of events somewhere and eliminate the latent unknown errors. Because *event trees* are used mainly as descriptions of consequences of an incident/accident, they can be used to find ways to mitigate or eliminate some negative outcomes, should an accident occur. This can be done through stopping the path of errors and eliminating latent failures.

The AEB model illustrates the *Barrier Function Failure Model* approach. According to barrier system function approaches, an accident is modeled as a number of errors in one sequential main trajectory of faults going through a number of failing defenses (Reason, 1990), called barrier functions by Svenson (1991). Barrier functions (e.g., a warning light, a blockage of a pathway) are executed by barrier function systems (e.g., a computer, a lead wall). Thus, before each error there is the possibility that a barrier function could have arrested the accident evolution. Using the barrier function failure approach, actions to eliminate or decrease the risk of incidents/accidents include elimination of the contributing errors in the accident evolution sequence and strengthening the barrier functions. The strengthening of those barrier functions is often effective in decreasing the risks of other accidents as well. Recently, Hollnagel (1999) has explored the barrier system and barrier function concepts and suggested a system for classifying different barrier systems and barrier functions.

CONCLUDING REMARKS

There are different ways in which an accident can be viewed an analyzed. Depending on the purpose, a particular perspective will be chosen, and different factors will be highlighted as explanations of an accident. To illustrate, a legal purpose has the goal of finding persons who can be attributed responsibility for an accident. Therefore, particular accident models are favored in which individuals causing the accident can be identified. Other perspectives include those originating from engineering, human factors, organizational and economic expertise (Svenson, Lekberg & Hammarberg, 1999). The role played by the perspectives, explicit and implicit goals of a method and an analyst may not always be obvious to those involved in performing and interpreting an analysis.Incidents and accidents represent sampling on the dependent variable and therefore it should always be kept in mind that good explanations do not necessarily lead to good predictions. In particular, the relative importance of different factors contributing to an accident, including those missing from an analysis, cannot be determined. Therefore, accident analyses alone cannot be used for determining an optimal set of actions to avoid similar accidents in the future. However, incidents and accidents provide feedback that something was wrong. Therefore the analyst should find out what went wrong and relate his or her findings to existing models of the systems that were involved. Based on a critical understanding of the accident analysis and systems knowledge the analyst should be able to suggest actions that can improve the safety of the systems analyzed.

REFERENCES

Dawes, R. M. (1993) Prediction the future versus understanding the past: A basic asymmetry. American Journal of Psychology, 106, 1 - 24.

Hale, A., Wilpert, B. & Freitag, M. (1997) After the event: From accident to organizational learning. Oxford: Pergamon.

Hollnagel, E. (1999) Accident analysis and barrier functions. Institutt for Energiforskning, Halden, Norway. IFE/HR/F-99/1121

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

Leveson, N. G. (1995) Safeware: System safety and computers. .New York: Addison-Wesley.

Reason, J. T. (1990) Human error. Cambridge UK , Cambridge University Press.

van der Schaaf, T. W. (1992) Near miss reporting in the chemical process industry. Ph D thesis, Eindhoven University of Technology.

Svenson, O. (1991) The accident evolution and barrier function (AEB) model applied to incident analysis in the processing industries.

Svenson, O. (1999) Learning from incidents and accidents through accident and barrier function (AEB) analyses. Manuscript.

Svenson, O., Lekberg, A. & Johansson, A.E.L. (1999) On perspective, expertise and differences in accident analyses: Arguments for a multidisciplinary integrated approach. Ergonomics, in press.

Acknowledgements

This study was supported by Netherlands Institute for Advanced Study in the Humanities and Social Sciences, the Swedish Nuclear Power Inspectorate and the International Atomic Energy Agency, Vienna. The views expressed in the paper do not necessarily reflect those of these organizations. The author is grateful for Erik Hollnagel´s highly significant and very helpful comments on an earlier version of this paper.

RÉSUMÉ

À PROPOS DES MODÈLES D'INCIDENTS ET D'ACCIDENTS

Un accident peut être expliqué de différentes façons selon le modèle d'analyse d'accident qui est utilisé. Quelques modèles fondamentaux d'accident sont présentés ici. Ils peuvent être appliqués isolément ou en combinaison les uns avec les autres. À ces modèles génériques sont associées différentes recommandations pour accroître la sécurité d'un système. Si un analyste ou une autorité qui souscrit à l'analyse ont une idée claire des modèles fondamentaux et des perspectives sous-jacentes à une analyse d'accident, leur évaluation des conclusions sera plus perspicace et moins biaisée que s'ils n'ont pas cette conscience. De plus, expliquer un accident n'est pas la même chose que le prédire et la cognition humaine ne fait pas aisément cette distinction. Par conséquent, le risque de tirer des conclusions causales impropres à partir d'analyses d'accidents peut tout aussi bien provenir des modèles fondamentaux d'accident que de la confusion entre explication et prédiction nécessaire pour adopter des mesures efficaces en vue d'éviter des accidents futurs.

MOTS CLÉS : Accident, Incident, Analyse d'accident, Analyse de risque, Facteurs humains.

ACCIDENTS AND BARRIERS

Erik Hollnagel
Graduate School of Human-Machine Interaction
University of Linköping, LIU/IKP/HMI, S-581 83 Linköping, Sweden
eriho@ikp.liu.se

Abstract

This paper discusses the barrier concept starting from a basic distinction between barrier functions, defined as the specific manner by which the barrier achieves its purpose, and barrier systems, defined as the organisational and/or physical foundation for the barrier function. Four different types are proposed, called material, functional, symbolic, and immaterial barrier systems respectively. A basic distinction between barrier functions is whether they are preventive or protective. This reflects whether the barrier function is intended to work before the occurrence of an accident or after it has happened. It is furthermore possible to describe a number of generic barrier functions, such as: containing, restraining, keeping together, dissipating, preventing, hindering, regulating, indicating, permitting, communicating, monitoring, and prescribing. There is no simple one-to-one correspondence between barrier functions and barrier systems, nor between barrier functions and their use as either preventive or protective barriers. The paper also introduces the specific discussion of the retrospective and prospective use of barriers.

Keywords

Accidents, failures, barriers, prevention, design, organisations.

INTRODUCTION

Accidents are frequently characterised either in terms of the events and conditions that led to the final outcome or in terms of the barriers that have failed. A barrier, in this sense, is an obstacle, an obstruction, or a hindrance that may either (1) prevent an action from being carried out or an event from taking place, or (2) prevent or lessen the impact of the consequences, for instance by slowing down the uncontrolled release of matter and energy, limiting the reach of the consequences or weakening them in other ways, cf. Figure 1. Barriers are important for the understanding and prevention of accidents. Firstly, the very fact that an accident has taken place means that one or more barriers have failed – i.e., that they did not serve their purpose or that they were missing. Secondly, once the aetiology of an accident has been determined and the causal pathways identified, barriers can be used as a means to prevent that the same, or a similar, accident takes place in the future.

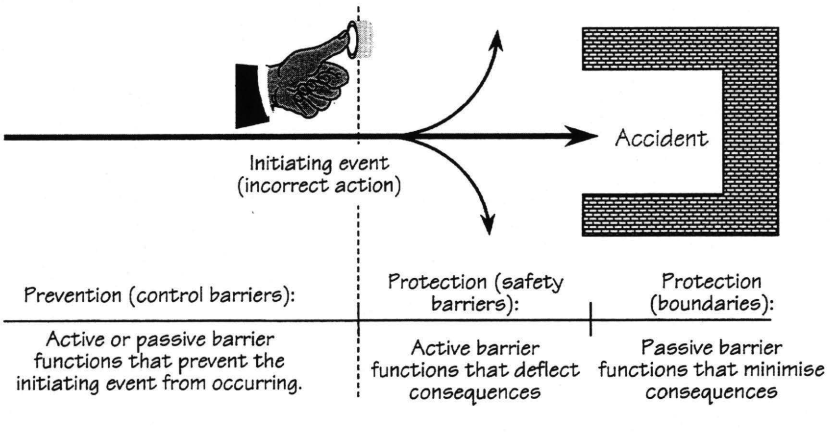


Figure 1: Use of barriers.

The notion of a barrier can be considered both in relation to a method or a set of guidelines for identifying barriers and in relation to a way of systematically describing or classifying barriers. The two aspects are, of course, not independent, since the method for analysis necessarily must refer to a classification scheme, regardless of whether the analysis is a retroactive or a proactive one (Hollnagel, 1998). As a starting point, a **barrier function** can be defined as the specific manner by which the barrier achieves its purpose, whereas a **barrier system** can be defined as the substratum or foundation for the barrier function, i.e., the organisational and/or physical structure without which the barrier function could not be accomplished. The use of the barrier concept should be based on a systematic description of various types of barrier systems and barrier functions, for instance as a classification system. This will help to identify specific barrier systems and barrier functions and to understand the role of barriers, in either meaning, in the history of an accident.

Despite the importance of the barrier concept, the accident literature only contains a small number of studies (Kecklund et al. 1996; Leveson, 1995; Svenson, 1991 & 1997; Taylor, 1998 and Trost & Nertney, 1985). The classifications proposed by these studies have been quite diverse, partly because of the lack of a common conceptual background, and partly because they have been developed for specific purposes within quite diverse fields. The most successful attempt of developing a theory of barriers has been the work of Svenson (1991), which also was the basis for the field studies of Kecklund et al (1996).

DESCRIPTORS OF BARRIER SYSTEMS

An analytical description of barriers can be based on several different concepts, such as the barriers' origin, their purpose, their location, and their nature. Of these, only the concept of the barrier nature is rich enough to support an extensive classification. The nature of barriers is principally independent of their origin, their purpose (e.g., as preventive or protective), and their location. By their nature barrier systems can range from physical hindrances (walls, cages) to ethereal rules and laws. A classification of barrier systems can be based on the following four main categories.

- **Material barriers** physically prevent an action from being carried out or the consequences from spreading. Examples of material barriers are buildings, walls, fences, railings, bars, cages, gates, etc. A material barrier presents an actual physical hindrance for the action or event in question and although it may not prevent it under all circumstances, it will at least slow it down or delay it. Furthermore, a material barrier does not have to be perceived or interpreted by the acting agent in order to serve its purpose.

- **Functional (active or dynamic) barriers** work by impeding the action to be carried out, for instance by establishing a logical or temporal interlock. A functional barrier effectively sets up one or more pre-conditions that have to be met before something can happen. These pre-conditions need not be interpreted by a human, but may be interrogated or sensed by the system itself. Functional barriers are therefore not always visible or discernible, although their presence often is indicated to the user in one way or another and may require one or more actions to be overcome. A lock, for instance, is a functional barrier, whether it is a physical lock that requires the use of a key or a logical lock that requires some kind of password or identification.

- **Symbolic barriers** require an act of interpretation in order to achieve their purpose, hence an "intelligent" agent that can react or respond to the barrier. Whereas a functional barrier works by establishing an actual pre-condition that must be met by the system, or the user, before further actions can be carried out, a symbolic barrier indicates a limitation on performance that may be disregarded or neglected. Alternative terms may therefore be conceptual or perceptual barriers. While the railing along a road is both a physical and a symbolic barrier, the reflective posts or markers are only a symbolic barrier: they indicate where the edge of the road is, but unlike the railing they are insufficient to prevent a car from going off the road. All kinds of signs and signals are symbolic barriers, specifically visual and auditory signals. The same goes for warnings (texts, symbols, sounds), interface layout, information presented on the interface, visual demarcations, etc.

- **Immaterial barriers** are not physically present or represented in the situation, but depend on the knowledge of the user to achieve their purpose. Immaterial barriers are usually also represented in a physical form such as a book or a memorandum, but are normally not physically present when their use is mandated. Typical immaterial barriers are: rules, guidelines, restrictions, and laws. In industrial contexts, immaterial barriers are largely synonymous with organisational barriers, i.e., rules for actions that are imposed by the organisation, rather than being physically, functionally or symbolically present in the system.

It is clearly possible to realise several barrier systems and functions in the same physical artefact or object. For instance, a door may have on it a written warning and may require a key to be opened. Here the door is a physical barrier system, the written warning is a symbolic barrier system, and the lock requiring a key is a functional barrier system. It may, in fact, be the rule rather than the exception that more than one type of barrier is combined, at least for the first three categories.

A Classification Of Barriers

The following Table 1, presents a classification of the barriers that are known from the general literature. Each barrier is described with regard to its **system**, i.e., one of the four main classes as defined above, and its **function** (or **mode**), i.e., the more specific nature of the barrier. The list of barriers presented here is clearly not exhaustive, but hopefully sufficiently extensive to be of some practical use.

Table 1: Barrier systems and barrier functions.

Barrier system	**Barrier function**	**Example**
Material, physical	***Containing*** or protecting. Physical obstacle, either to prevent transporting something **from** the present location (e.g., release) or **into** present location (penetration).	Walls, doors, buildings, restricted physical access, railings, fences, filters, containers, tanks, valves, rectifiers, etc.
	Restraining or preventing movement or transportation.	Safety belts, harnesses, fences, cages, restricted physical movements, spatial distance (gulfs, gaps), etc.
	Keeping together. Cohesion, resilience, indestructibility	Components that do not break or fracture easily, e.g. safety glass.
	Dissipating energy, protecting, quenching, extinguishing	Air bags, crumble zones, sprinklers, scrubbers, filters, etc.
Functional	***Preventing*** movement or action (*mechanical, hard*)	Locks, equipment alignment, physical interlocking, equipment match, brakes, etc.
	Preventing movement or action (*logical, soft*)	Passwords, entry codes, action sequences, pre-conditions, physiological matching (iris, fingerprint, alcohol level), etc.
	Hindering or impeding actions (spatio-temporal)	Distance (too far for a single person to reach), persistence (dead-man-button), delays, synchronisation, etc.
Symbolic	***Countering,*** preventing or thwarting actions (visual, tactile interface design)	Coding of functions (colour, shape, spatial layout), demarcations, labels & warnings (static), etc. *Facilitating correct actions may be as effective as countering incorrect actions.*
	Regulating actions	Instructions, procedures, precautions / conditions, dialogues, etc.
	Indicating system status or condition (signs, signals and symbols)	Signs (e.g., traffic signs), signals (visual, auditory), warnings, alarms, etc.
	Permission or authorisation (or the lack thereof)	Work permit, work order.
	Communication, interpersonal dependency	Clearance, approval, (on-line or off-line), in the sense that the lack of clearance etc., is a barrier.
Immaterial	***Monitoring***, supervision	Check (by oneself or another a.k.a. visual inspection), checklists, alarms (dynamic), etc.
	Prescribing: rules, laws, guidelines, prohibitions	Rules, restrictions, laws (all either conditional or unconditional), ethics, etc.

It is not always easy to classify barriers. A wall is, of course, a physical barrier and a law is an immaterial barrier. But what about a procedure? This by itself is an instruction for how to do something, hence not primarily a barrier. Procedures may, however, include warnings and cautions, as well as conditional actions. The procedure may exist as a physical document, but it works because of its contents or meaning rather than because of its physical characteristics. The warnings, cautions, and conditions of a procedure are therefore classified as a symbolic barrier system, i.e., they require an act of interpretation in order to work.

Symbolic barriers often complement immaterial barriers. For instance, road signs supplement the general speed limits given by the traffic laws. Symbolic barriers may also complement material barriers to encourage their use. Seat belts are material barriers, but can only serve their purpose when they are used. In commercial aircraft, seat belt use is supported by both static cautions and dynamic signals (seat belt sign), as well as a visual inspection. In private cars the material barrier is only supported by the immaterial barrier, i.e., the traffic laws, which often produces a less than satisfactory result.

BARRIERS, ACCIDENT ANALYSIS, AND SYSTEM DESIGN

In order for a classification to be useful, it must be closely integrated with a method. In the case of barriers, there is a need of two sets of different methods, one considering the identification of barriers in accident analysis, and the other the specifications of barriers for system design.

For the purpose of an accident analysis, barrier identification is generally carried out in a rather *ad hoc* fashion. The common practice in risk analysis is to look for known barriers - similar to the search for latent failure conditions, sneak paths, or failure modes - and this approach has simply been applied to accident analysis as well. The principal disadvantage is that the barrier analysis in this way is carried out on its own, rather than as an integral part of the general accident analysis method. Although risk analysis has some similarities to accident analysis, it is clearly not a complete accident analysis method by itself, since it does not address aspects such as accounting for the interaction between the various elements of the socio-technical system, or describing the common performance conditions. It is therefore necessary to find a way of incorporating a systematic classification of barriers into common accident analysis methods. The easiest solution is presumably to combine the generic fault tree analysis with a barrier analysis, to identify the risks emanating from the failure of barriers, which can be described as input conditions to the logical gates.

For the purpose of system design, the emphasis is on how to ensure that the system functions as specified. While this clearly is an essential achievement, it is also important to consider how the system may not function as specified, i.e., how it may fail. Such analysis are common in the case of complex technological systems, e.g. as fault trees, cause-consequence analyses, event trees, FMEA, HAZOP, etc., but are conspicuous by their absence in the case of interactive systems - perhaps with the notable exception of HRA. It is, however, of the utmost importance to use barriers as a pivotal element in system design, since it is only by a inventive combination of barriers and facilitators that an effective and safe system functioning can be achieved.

For event trees, barriers are uncomplicated to insert since they are represented simply as failures, or rather - effective barriers are represented in terms of successes or very low failure probabilities. It is then up to the designer later on to be more specific about the types of barriers that may be needed to achieve the desired probability.In that sense there is a gradual transition to cause-consequence trees, which are more developed in the forward direction than event trees. Here the introduction of the logical gates means that barriers become more tangible and must be specified in greater detail.

Since barriers are included in a system to prevent undesirable events from occurring or to protect against their consequences, it is important that potential barrier failures themselves can be assessed, so that the weaknesses of the system are known. A tentative description of the conditions that are required for adequate barrier functioning is shown in Table 2.

Table 2: Requirements for effective barrier functions.

Barrier system	Barrier function	Pre/condition for proper functioning
Material	Physical.	Reliable construction, possibly regular maintenance.
Functional	Mechanical	Reliable construction, regular maintenance.
Functional	Logical	Verified implementation, adequate security.
Functional	Spatio-temporal	Reliable construction, regular maintenance.
Functional	Monitoring	Reliable performance of monitor
Symbolic	Interface design	Valid design specification, verified implementation, systematic updating
Symbolic	Information	High-quality interface design, reliable functioning.
Symbolic	Signs, signals and symbols	Regular maintenance, systematic modification,
Symbolic	Lack of permission or authorisation	High compliance by users.
Immaterial	Communicative, interpersonal	Nominal working conditions (no stress, noise, distraction, etc.).
Immaterial	Rules, cautions, warnings, prohibitions	High compliance by users.

In order to include the concept of barriers in accident analysis and accident prevention, it is necessary to combine the barrier concept with the notion of error modes. Hollnagel (1998) identified eight basic error modes for human actions, which later were extended to cover systemic failure modes as shown in Table 3 (cf. Hollnagel, 1999).

Table 3: Human and systemic error modes.

	Human error mode	Systemic error mode
Timing	Action performed too early or too late	Position reached too early or too late. Equipment not working as required.
Duration	Action performed too briefly or for too long	Function performed too briefly or for too long. System state achieved too briefly or held for too long
Distance	Object/control moved too short or too far	System or object transported too short or too far
Speed	Action performed too slowly or too fast	System moving too slowly or too fast Equipment not working as required.
Direction	Action performed in the wrong direction	System or object (mass) moving in the wrong direction
Force / power / pressure	Action performed with too little or too much force.	System exerting too little or too much force. Equipment not working as required.
		System or component having too little or too much pressure or power.
Object	Action performed on wrong object	Function targeted at wrong object
Sequence	Two or more actions performed in the wrong order,	Two or more functions performed in the wrong order,
Quantity and volume	None	System/object contains too little or too much or is too light or too heavy.

In order to be able to select the right barrier during system design, it is necessary to assess the efficiency of each barrier system relative to the failure or error modes. Consider, for instance, the error mode of distance. Here a material barrier can be highly efficient in preventing a movement from being taken too far (although not for preventing too short a movement). A functional barrier may also be highly efficient, but both symbolic and immaterial barriers are likely to be of little use.

The analyses made so far have indicated that immaterial barriers normally are rather inefficient, even though they are cheap and fast to implement. This corresponds to the ordering of approaches to hazard elimination in the MORT technique (Knox & Eicher, 1983), where immaterial barriers, such as the development of special procedures to handle the situation, come last. The other barrier systems may be efficient in different ways, and in practice the establishing of an effective barrier requires a combination of several barrier systems. Guidelines and principles for how this is to be done will be developed in a recently started project.

REFERENCES

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

Hollnagel, E. (1999). *Accident analysis and barrier functions.* Halden, Norway: Institute for Energy Technology.

Kecklund, L. J., Edland, A, Wedin, P. & Svenson, O. (1996). Safety barrier function analysis in a process industry: A nuclear power application. *Industrial Ergonomics, 17,* 275-284.

Knox, N. W. & Eicher, R. W. (1983) *MORT user's manual* (DOE 76/45-4). Idaho Falls, Idaho: EG&G Idaho, Inc.

Leveson, N. (1995). *Safeware. System safety and computers.* Reading, MA: Addison-Wesley Publishing Company.

Svenson, O. (1991). The accident evolution and barrier function (AEB) model applied to incident analysis in the processing industries. *Risk Analysis, 11*(3), 499-507.

Svenson, O. (1997). *Safety barrier function analysis for evaluation of new systems uin a process industry: How can expert judgment be used?* In: Proceedings of Society for Risk Analysis Europe Conference, Stockholm, June 15-18, 1997.

Taylor, R. J. (1988). *Analysemetoder til vurdering af våbensikkerhed.* Glumsø, DK: Institute for Technical Systems Analysis.

Trost, W. A. & Nertney, R. J. (1985). *Barrier analysis* (DOE 76-45/29). Idaho Falls, Idaho: EG&G Idaho, Inc.

RÉSUMÉ

ACCIDENTS ET BARRIÈRES

Cette communication est une discussion du concept de barrière qui s'appuie sur une distinction fondamentale entre les fonctions de barrière, définies comme la façon dont une barrière atteint son but, et les systèmes de barrière, définis comme le fondement organisationnel et/ou physique d'une fonction de barrière. Quatre types de systèmes de barrière sont proposés : matériels, fonctionnels, symboliques et immatériels. On distingue des fonctions de barrière préventives et protectrices, selon qu'une barrière intervient avant ou après l'accident. Il est alors possible de décrire un certain nombre de fonctions de barrière génériques, telles que contenir, restreindre, maintenir ensemble, dissiper, prévenir, empêcher, réguler, indiquer, permettre, communiquer, surveiller et prescrire. Il n'y a de correspondance univoque simple, ni entre les fonctions de barrière et les systèmes de barrière, ni entre ces fonctions et leur caractère préventif ou protecteur. La communication traite également de l'utilisation rétrospective et prospective des barrières.

MOTS CLÉS : Accidents, Défauts, Barrières, Prévention, Conception, Organisations

Session 7: Complex and dynamic systems
Les systèmes complexes et dynamiques

Timing and its relation to strategies in controlling dynamic systems

Eva Carling

Department of Psychology, Uppsala University,
Box 1225, S - 751 42 Uppsala, Sweden
ecg@hig.se

ABSTRACT

This paper describes experiments within the microworld paradigm and how they can cover dynamic aspects of human decision making. It discusses if results from experiments made in different microworlds can support ideas that timing is an important factor in dynamic decisionmaking. It also suggests that timing is a complement to strategies while performing well in complex systems.

Keywords

Complex systems, dynamic decision making, microworld, strategies, timing.

INTRODUCTION

The task of process control is an example of dynamic decisionmaking. This kind of decisionmaking can be described by saying: It is constituted of a series of decisions (i), which are interdependent (ii), causing changes that are a result both of the decision maker's interventions and of the system's internal dynamics (iii) and which have to be made in real time (iv).

In the research carried out within the framework of dynamic decisionmaking (see Brehmer& Allard, 1991, Dörner & Schaub, 1994, Sterman, 1989), ecological models (Amalberti, 1998, Rasmussen, 1993, Vicente, 1996), and complex problem solving (Frensch & Funke, 1995) we have learnt that some characteristics of the situation are critical in controlling a process. Complexity can be defined by a) many variables and side effects, b) lack of transparency (i.e. not visible effects that necessitates inferences), and c) dynamics resulting in delay of feedback. Complexity is a matter that humans find hard to cope with, whether one considers one of its characteristics at a time or in combination. We also know that the temporal aspects of feedback are hard to learn, especially when it is necessary to take on further actions when the feedback on the earlier ones is delayed. We have seen in research from different areas as process control (Crossman & Cooke, 1974, replicated by Moray, Lootstein & Pajak, 1986; Hoc, 1989), enterprise stocks and flows (Sterman, 1989) and fire fighting (Brehmer and Allard, 1991; Omodei, Anastasios, Waddell, & Wearing, 1995) the limitations of humans while controlling systems with feedback delays.

In this paper I am describing some results from studies carried out by our group in Uppsala. These results point out a connection between a cognitive strategy, based on a mental model of the system or the process and a more effortless, but not automatic control strategy, characterised by direct feedback. That the latter mentioned strategy really follows after a cognitive one is documented in some of the participant's behaviour, and is in other cases indirectly shown in their answers when queered about the causality in the system. I will argue that the connection between the strategies is related to the concept of timing, which besides the mental model also is an important success factor in controlling dynamic processes. Very little is known about how time is handled in decisionmaking, timing is mostly studied in motor actions and in music perception and production (Jones, 1990). Timing in decisionmaking may be described as being when someone attends to the states and the history of the processes involved (Michon, 1990) and then allocates resources or takes actions in a preferred order and/or in a certain interval of the processes.

COMPARING ENGINEERING STUDENTS AND PROCESS OPERATORS

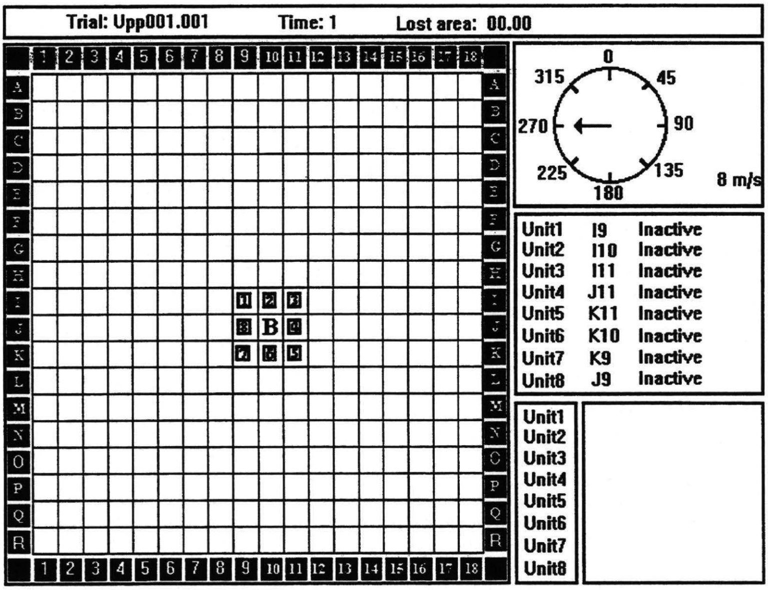
In a study (Carling, 1998) was shown how participants with theoretical and practical experience of process control learnt to control a simulated evaporator, i.e. a process to take away liquid from a solution, in the simulated technique used, by heating the solvent under pressure. A series of tasks were constructed with the simulated evaporator, aiming at covering many different difficulties, like coping with disturbances, moving the process to a balance at a new level and pushing the system towards its constraints. After having gone through three learning sessions with the simulated evaporator, a fourth session was administered to validate the data from the last learning sessions. The subjects studied consisted of two groups, one with four students who had taken grades in process control, and another with four experienced operators from a paper mill. They were told their goal was to avoid that the evaporator would get empty or to overflow, that it would explode or implode and was in some of the tasks asked to fulfil the goal to get an outflow of a certain concentration. After the trials questions concerning the causal relations in the system were answered by both groups of participants.

Several measures of performance were calculated in order to describe the development of a mental model needed for control. Some of them measured systems loss measured as deviation from the balance level finally reached, per time unit used, some included amount of control measures used, and some tried to consider the dependency, that one control is conditioned on the result of the earlier made control, as well as signs of learning via feedback to stop making errors. The understanding of the causal relations in the system was good in both groups. A principal component analysis gave four preliminary components which could be thought of as describing the mental model used in the tasks. One was the *Factor of system loss*, a second one was a *Right-right-factor* (series of correct control movements), the third a *Correcting-ability factor* and finally the *Control-*

time factor. The results showed that in the last of the two learning sessions the participants with practical experience made more correct control moves and could correct wrong moves better than those participants without practical experience. A participant with theoretical background controlled with lesser system loss and shorter time per control than those without. In the fourth, validating session the difference between the groups was not stable, indicating that the learning sessions had not been long enough. However, the way the patterns changed was interesting, as it showed that the 'theoretical' participants turned into the strategy shown by the practical ones. A performed discriminant analysis showed that the four factors classified all subjects correctly as Practical or Theoretical in the calibration (the last two learning sessions) but in the validating fourth session all the engineering students were classified as Practical!

DIFFERENT STRATEGIES IN A FIRE FIGHTING SIMULATION

In another study, the simulation used was NEWFIRE, where the participants were asked to act as a fire chief in charge of eight fire fighting units. From a spotter plane she/he got reports of the position and status of the fire fighting units (Löwborg & Brehemer, 1991).


Trial: Upp001.001 Time: 1 Lost area: 00.00

Unit1 I9 Inactive
Unit2 I10 Inactive
Unit3 I11 Inactive
Unit4 J11 Inactive
Unit5 K11 Inactive
Unit6 K10 Inactive
Unit7 K9 Inactive
Unit8 J9 Inactive

Figure 1. The interface of NEWFIRE at the start of a trial. The eight fire fighting units are surrounding the Base (B). The direction and strength of the wind is shown in the upper small window to the right.

The goal is to save the base (otherwise the simulation will come to an end) and to keep as much forest as possible out of the fire (or fires). The participants consider the information of where the fire fighting units are and of their status; if they are standing idle, mobilising or demobilising or busy fighting the fire. She/he thereby tries to allocate these fire fighting units so that they can extinguish the forest fires as soon as possible. In one of our studies (Carling & Rigas, 1999) the participants were asked to perform twenty fire fighting trials with different fire scenarios. In the analysis of their performance data we looked for both non-dynamic cognitive measures of strategy and measures of the timing of their actions with the fire fighting units. We thoroughly analysed four cases, where the participants developed two instances of successful and two of poor strategies.

What could be seen from these cases was that one of the successful strategies is characterised by feed-forward planning. This strategy presupposed both that the participant could predict he propagation of the fire, that she/he managed to estimate the time that would be needed before the units were in place, as well that it was necessary that the participant handled an efficient resource allocation. Use of this strategy leads to high memory load and cognitive effort, but it does not demand motor speed and attention. The second strategy is characterised by direct feedback control in order to minimise dead times (time before the units can be ready to start working) by continuously reallocating the units. This latter strategy does not cause any memory load, nor is does it demand extensive central processing or attention, rather it is using 'direct perception' or ad hoc decisionmaking.

TRYING TO JUDGE THE IMPACT OF TIMING IN THE FIRE FIGHTING SIMULATION

In a study of the effects of feedback delay in dynamic decisionmaking, I have been analysing the components of a successful mental model (Carling, 1997) but I did also consider the importance of timing for good performance. Even while taking into consideration that a model of the system is a precondition of being able to control the system, it is not enough to use the model as a set of rules. To use the rules at the right point of time or at least within a time interval is necessary for good performance. The simulation used here was again NEWFIRE.

The operational definition used in this study was that timing was the number of times within a trial when the participant allocated a unit into the fire instead of allocating it to the outskirts of the fire and the fire thereby spread outside the fire fighting unit. When she/he did this though the fire had been on for such an amount of time so that she/he should have judged the risk of spreading as big. A measure of timing would therefore not become a number, but rather an interval, within which the action could be taken. Within this interval pure speed of action would surely make an impact, but this is not captured by my measure used in this study.

The results showed significant correlation between timing and performance. What was interesting with regards to our line of reasoning here are the results showing that in this study, while the components, constituting the mental model, explained .25 of the variance in performance, the timing measure alone also explained as much as

.25. Another result was that timing correlated low with intelligence, as measured by APM (Advanced Progressive Matrices, prepared by J.C. Raven), to be compared with the results from another study where the mental model measure correlates somewhat higher with intelligence (Rigas, Carling, & Brehmer, 1998). In a recently performed study, (Friberg and Nilsson, 1999) these results were repeated and also appeared in another microworld.

The results therefore indicate that timing is a factor highly decisive for performance in controlling dynamic systems, and that it is a factor that is apart from the structural aspects of the mental model.

CONCLUSIONS

Considering that in the study on the evaporator, both the group who learnt the strategy via building a mental model (the engineering students) and the group who already had an internalised model (the operators) were able to answer the questions on the causal relations in the systems correctly.

In the NEWFIRE study however, the case studies indicate that individual differences may be a factor in explaining the choice of strategy, as not all of the participants changed to a more 'relaxed' or direct feedback-governed mode in contrast to the first study where all participants changed strategy. The reason may be that the engineering students of the first study formed a more homogenous group than the one in the second study.

Timing competence, as referred to in the last studies, seems to correlate low with intelligence, and a relevant hypothesis to test is that timing goes better together with the relaxed mode, with rhythm or at least ordering or counting. Yet it may well be that timing also is exerted in other, perhaps more cognitive ways, which are more readily included in a mental model. Alternatively, switching to a more ad hoc oriented strategy liberates resources needed in these presumably more cognitively demanding timing behaviour. However, both the composite measure of mental model and the timing measure need to be much refined before we can tell if they are orthogonal, or else, what is needed to capture dynamics in a mental model theory. There is certainly a need for further study of 'Cognitive models of timing in decisionmaking' to paraphrase Block (1995).

REFERENCES

Amalberti, R. (1998). *Why Operators' Cognitive Models are Hard to Incorporate into Design? The Case of Human Reliability Models.* Keynote paper presented at the Second European Conference on Cognitive Modelling, Nottingham, UK.

Block, R.A. (1995). *Cognitive Models of Psychological Time.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Brehmer, B. & Allard, R. (1991). Real-time, dynamic decision making. In Hogarth (Ed.), *Insights in decision making.* Chicago, IL: University of Chicago Press.

Carling, E. (1997). *Timing in Dynamic Decision Making.* Paper presented to the sixteenth *Conference on Subjective Probability, Utility, and Decision Making.* Leeds, England.

Carling, E. (1998). *Learning a mental model of a simulated evaporator - the impact of background knowledge.* Manuscript in preparation.

Carling, E. & Rigas, G. (1999). Strategies in a Real Time Dynamic Decision Making Task. In N. Georgantzis (Ed.) *Spatial Economics and Ecosystems* Southampton: WIT Press/ Computational Mechanics Publications, in press.

Crossman , E.R.F.W., & Cooke, J.E. (1974). Manual control of slow response systems. In E. Edwards & F.P. Lees (Eds.). *The human factor in process control.* London: Taylor & Francis.

Dörner,D.& Schaub,H.(1994). Errors in Planning and Decision Making and the Nature of Human Information Processing. *Applied Psychology: An International Review Special Issue on Human Error; Memorandum Lst Psychologie II Universität Bamberg*, 2,10.

Frensch, P.A., & Funke, J. (Eds.) (1995). *Complex problem Solving. The European Perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Friberg, M. & Nilsson, E. (1999). *Important factors in decisionmaking in complex systems - a study in two computer simulated microworlds.* C-level thesis, Department of Mathematics, Natural and Computer Science, University of Gavle. (in Swedish).

Hoc, J.M. (1989). Strategies in controlling a continuous process with long response latencies: Needs for computer support to diagnosis. *International Journal of Man-Machine Studies,* 30, 47-67.

Jones, M.R. (1990). Musical Events and Models of Musical Time. In R.A. Block *Cognitive Models of Psychological Time* Hillsdale, NJ: Lawrence Erlbaum Associates.

Lövborg, L., & Brehmer, B. (1991). *NEWFIRE- A Flexible System for Running Simulated Fire-Fighting Experiments.* Risö National Laboratory, Roskilde, Denmark.

Michon, J.A.(1990). Implicit and Explicit Representations of Time. In R.A. Block *Cognitive Models of Psychological Time* Hillsdale, NJ: Lawrence Erlbaum Associates.

Moray, N., Lootstein, P., & Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-16, 497-504.*

Omodei, M. M., Anastasios, J., Waddell, F., & Wearing, A. J. (1995*). Person Characteristics in Dynamic Decision Making.* Paper presented at SPUDM-15, 1997, Jerusalem, Israel.

Rasmussen, J. (1993) Diagnostic reasoning in action. *IEEE transactions in Systems, Man and Cybernetics., SMC.*

Rigas, G., Carling, E., & Brehmer, B. (1998) *Do the close to zero correlations between test-intelligence and dynamic decision making performance depend on the static character of intelligence tests?* Intelligence, submitted.

Sterman, J.D. (1989). Mispercceptions of feedback in dynamic decisionmaking. *Organizational Behavior and Human Decision Processes, 43, 301-355.*

Vicente, K.J. (1996). Improving dynamic decision making in complex systems through ecological interface design: A research overview. *System Dynamics Review,* 12, 251-279.

RÉSUMÉ

LA SYNCHRONISATION ET SES RELATIONS AUX STRATÉGIES DANS LE CONTRÔLE DES SYSTÈMES DYNAMIQUES

Cette communication décrit des expériences utilisant le paradigme du micromonde et la façon dont elles peuvent couvrir les aspects dynamiques de la prise de décision. Elle discute du caractère probant des résultats d'expériences réalisées sur différents micromondes en ce qui concerne l'importance de la synchronisation dans la prise de décision dynamique. Elle suggère aussi que la synchronisation est un bon complément aux stratégies efficaces dans les systèmes complexes.

MOTS CLÉS : Systèmes complexes, Prise de décision dynamique, Micromonde, Stratégies, Synchronisation.

Context-based reasoning and decision graphs. Application in incident management on a subway line.

Patrick Brézillon, Laurent Pasquier & Ilham Saker

Laboratoire d'Informatique de Paris 6, Case 169, Université Paris 6
4, place Jussieu, 75252 Paris Cedex 05, France
Tel: 33 1 44 27 70 08 - Fax: 33 1 44 27 70 00
E-mail: Forename.Name@lip6.fr

ABSTRACT

Subway line control is a difficult task and context plays an important role in the way incidents are managed. The modelling of procedures and practices for incident solving is usually made by decision trees, but taking the dynamism of contextual states (contextual knowledge or proceduralised context) into account implies to extend the model. Times notions (such as precedence and independence) are also included in this decision graph model.

Keywords

Decision tree, Decision graph, Context, Subway control.

INTRODUCTION

The management of incidents on a line of subway is a difficult task for several reasons. On the one hand, the overall strategy is to retrieve rapidly a normal service on a complex and dynamic system. This global strategy constrains all the incident solvings. On the other hand, parameters that intervene when an incident occurs are numerous and of various natures. The SART project (SART is the French acronym for support system in traffic control, and general information may be found at http://www.lip6.fr/SART) aims at the design and development of an intelligent system to support the operator who is responsible for a subway line at the RATP (the French company of subway in Paris) when an incident occurs (Brézillon, 1997).

SART is a multiagent system with now three agents (Brézillon, 1998): a line configurator, a traffic simulator and an incident manager. The incident manager, which is the subject of this paper, is planned to support the operator in the solving of incidents by recalling him past similar incidents and suggesting strategies for solving the incident at hand.

Incident solvings and their contexts can have different natures and most of them are the object of official procedures of resolution that are based on the experience acquired by the RATP. However, strategies selected by operators arise from a compromise between efficiency and rapidity of the normal-control retrieval. As a result, the strategy to solve a complex incident is different from one operator to another because the choice of a strategy upon another one to solve an incident depends on the context in which the incident occurs.

Thus, one must (i) take context into account in the models of operators' reasoning, (ii) differentiate between procedures and strategies effectively used, and (iii) deal with the personality and subjectivity of operators. The notion of context plays an important role in the SART project, as, indeed, in numerous domains (Brézillon, 1999). For example, a problem of traction in a train will be solved differently according to the fact that the incident occurs during rush hours or not.

THE CONTEXTUAL DIMENSION OF THE PROBLEM

Our definition of context is the set of elements that do not intervene directly in the incident solving but constrain it. Contextual elements are either known beforehand and can then be acquired automatically (e.g., position of trains on the line), or acquired from operators during the incident solving (e.g., the number of travellers at a given time). Brézillon & Pomerol (1999) define contextual knowledge as all the knowledge that is relevant for one person in a given situated decision problem and can be mobilised to understand that problem. Contextual knowledge is evoked by situations and events, and loosely tied to a task or a goal. Although the contextual knowledge exists in theory, it is actually implicit and latent, and is not usable unless a goal (or an intention) emerges. Contextual knowledge is a part of the "context", the rest of the context, which is not relevant for the situation, is called external knowledge.

When an event occurs, the attention of the actor is focused on it and a large part of the contextual knowledge is proceduralized. We call the proceduralized part of the contextual knowledge, at a given step of a decision making, the proceduralized context. The proceduralized context is invoked, structured and situated according to a given focus. It is like "compiled" knowledge and is generally elicited with the usual techniques of knowledge acquisition.

At a given step of a decision making, one has: proceduralized context that is knowledge commonly known by the actors of the problem and directly (but tacitly) used for the problem solving; contextual knowledge that is knowledge not explicitly used but influencing the problem solving; and external knowledge that is knowledge having nothing to do with the current decision making step, but known by many actors of the problem.

There are some similar views in the literature. For example, Anderson's theories (1993) assume that knowledge is first acquired in a declarative form which encodes the basic facts and examples found in the instructions (our contextual knowledge). Once acquired, this knowledge is used by general problem-solving rules to create rules specific to a given context (our proceduralized context). Turner (1999) considers descriptive and prescriptive knowledge that are close from contextual knowledge and proceduralized context here. There is another parallel with the "global-context" and "local-context" in (Carenini, 1993). "Global-context" (the contextual knowledge here) indicates the current topic under discussion. It contains the place in the dialogue history where this topic was begun. "Local-context" (the proceduralized context here) points to the most recent utterance. As regards Schank's theory, assuming that context is the set of all the possible stories, the case corresponds approximately to our proceduralized context while contextual knowledge is the set of paradigmatic cases.

PROCEDURES AND PRACTICES

At the RATP, most of the incidents have been well-known for a long time (object on the track, lack of power supply, suicide, etc.). Thus, the company has established procedures for incident solving on the basis of their experience. However, each operator develops his own practice to solve an incident, and one observes almost as many practices as operators for a given procedure because each operator tailors the procedure in order to take into account the current proceduralised context, which is particular and specific. In many working processes human beings can be observed to develop genuine procedures to reach the efficiency that decision makers intended when designing the task. Some parts of this practice are not coded (Hatchuel, 1992). Such know-how is generally built up case by case and is complemented by "makeshift repairs" (or non-written rules) that allow the operational agents to reach the required efficiency. This is a way of getting the result whatever the path followed. The validation of those unwritten rules is linked more to the result than to the procedure to reach it. De Terssac (1992) spoke of logic of efficiency.

The modelling of operators' reasoning is a difficult task because operators use a number of contextual elements, and by the fact that procedures for solving complex incidents have some degree of freedom. Their reasoning stems from some chunks of implicit knowledge which are imposed on the driver because they correspond to mandatory procedures. Procedures are established from operator's experience during similar incidents and fixed by the company. As such, procedures are proceduralized contexts. An implicit piece of knowledge is that travellers are safer in a station than in a tunnel. At a deeper level, the driver has to avoid stopping the train a long time in a tunnel because some travellers may have behavioural troubles such as claustrophobia and could leave the train to wander about on the railway (and thus may generate another type of incident such as "Traveller on the railway"). These pieces of knowledge, which are not necessarily expressed, result in more or less proceduralized actions that are compiled as the proceduralized context. Very often many pieces of proceduralized context are structured together in comprehensive knowledge about actions.

Moreover, there is no procedure for complex incidents, but a set of procedures for solving parts of the incident. For example, when a train cannot move in a tunnel, there are procedures for evacuate travellers at the nearest station, for evacuate the damaged train by another train, etc. Some procedures are sequential, but others may be accomplish in any order compared to some ones. For example, when a train must push a damaged train, both trains must be empty but the order in which travellers of the two trains are evacuated is not important and depends mainly on the context in which are trains. What is important is that the two actions must be accomplished. As a consequence, there are as many strategies for solving an incident as operators: Cases that are similar in one context may be totally dissimilar in others as already quoted by Tversky (1977). This observation has strong consequences on the representation of operators' reasoning.

MODELING OF THE REASONING

Our modelling associates case-based reasoning and decision trees and includes an explicit representation of contextual knowledge (Pasquier, 1999). Case-based reasoning permits to manage a large number of cases and possesses a power of generalisation. Decision trees are less flexible but express finely operators' reasoning. A case contains the description of an incident, its contextual elements and the strategy used to solve it.

As most of the contextual elements may intervene in several scenarios (e.g., traffic activity, position of the next train), operators prefer to take them into account as soon as possible to get a general picture of the best path to choose. At this step, contextual knowledge is proceduralized and in the meantime operators postpone action. The main objective is to eliminate event nodes. By grouping together a set of actions in a macro-action, operators hope to make the following step easier. Figure 1 gives a simplified example of a decision tree drawn from the SART project (see Table 1 for explanations).

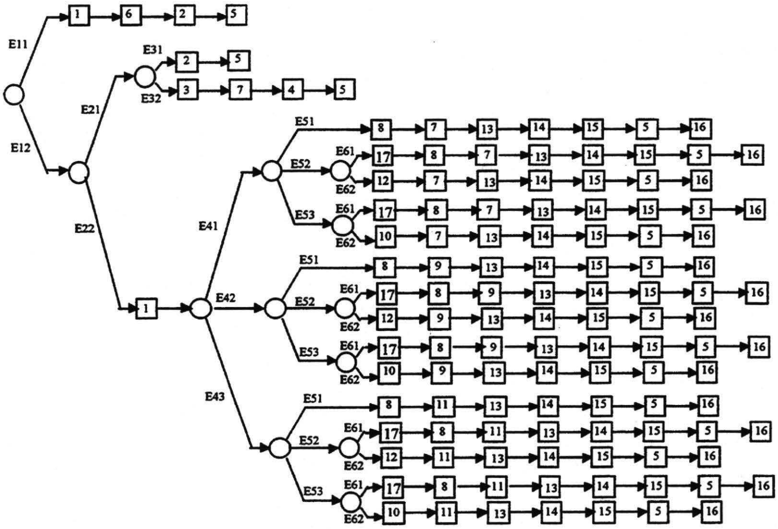


Figure 1: A tree in which many actions are postponed to the end of the branches (symbols are described in the following tables)

Actions	
1	Residual traffic regulation
2	Damaged train continues with travellers
3	Damaged train continues with travellers until a steep incline
4	Damaged train restarts without travellers
5	Park damaged train at end station
6	Repair damage
7	Exit of damaged train's travellers
8	Exit of next train's travellers
9	Exit of damaged train's travellers via available cars
10	Exit of next train's travellers via available cars
11	Exit of damaged train's travellers via track
12	Exit of next train's travellers via track
13	Next train joins damaged train
14	Link both trains
15	Convoy return to end station
16	Unlink damaged train
17	Next train goes to next station

Events	
E11	Immediate repair possible
E12	Immediate repair impossible
E21	Enough power units available
E22	Not enough power units available
E31	No steep incline between damaged train and end station
E32	Presence of steep incline until end station
E41	Damaged train at station
E42	Damaged train under tunnel
E43	Damaged train partially at station
E51	Next train at station
E52	Next train under tunnel
E53	Next train partially at station
E61	Presence of a station between damaged train and next train
E62	No station between both trains

Table 1: Some Actions (number) and events (E + number) in case of incident on a metro line

Without entering the details, macro-actions are a way to proceduralise contextual knowledge and to introduce modularity in the diagnosis process by managing different modules accomplishing the same function in different ways according to the context. However, action postponement is not always possible, and it is preferable to look for pruning the decision tree in some situations (Brézillon, 1998). On Figure 1, one can identify several macro-actions existing on different branches (e.g. A6-A7-A9). Such macro-actions are a kind of compilation, originated from experience, of several actions. In this compilation, a part of the knowledge on each action becomes implicit in the proceduralized context. This is close to Edmondson & Meech's view (1993) on context as a process of contextualisation. However, for explaining a macro-action (and thus the whole reasoning involved in), an operator needs to decompile the macro-action for retrieving the rationales. Such an operation is not always easy, especially when experience comes from previous generation of operators.

FROM DECISION TREES TO DECISION GRAPHS

In our application the goals of all branches of the decision tree are the same: "Fast retrieval of the normal control of the line" and branches express only different means to reach the final state. Thus, our goal is to find the best way at a given time and in a given context to retrieve a normal control.

Our representation also integrates the distinction between contextual and contextualised knowledge made in (Brézillon, 1998; Pomerol, 1999). Contextual knowledge constrains implicitly an incident-solving step while contextualised knowledge intervenes explicitly in it. However, the distinction must be considered as dynamic because the knowledge status (contextual or contextualised) changes from one step of the incident solving to the following one. This dynamism is modelled by joining back branches separated by an alternative on a proceduralised context as soon as this one return at a contextual state.

Operators often delay their decision making in order to collect before hand a maximum of contextual information. Thus, they constitute sequences of actions called macro-actions (Brézillon, 1998). This is a kind of diagnosis based on operators' experience (Pomerol, 1997). One interest of macro-actions is a better visibility of the decision tree because one object represents several other objects. Macro-actions are general and usable in different solvings of incidents.

The order in which some actions are executed may be indifferent. For example, when a train must push another train with a problem, both trains must be empty but the order in which the trains are made empty is not important and depends mainly on the context in which are trains. We thus have added temporal branching in the decision tree to the classical conditional branching (Pasquier, 1999). Temporal branchings represent branches that are independent at a temporal level.

Moreover, the decision making is at two levels: a local level that concerns the incident solving itself, and a global level that concerns the whole line and other trains. For example, the macro-action "Control" aims at minimise the effect of an incident on the whole line when the treatment of the incident is a parallel task. Indeed, with these new considerations, context dynamism management, macro-actions and temporal branching, we face more exactly a decision graph rather than a decision tree. Figure 2 gives the graph representation of the decision represented in Figure 1.

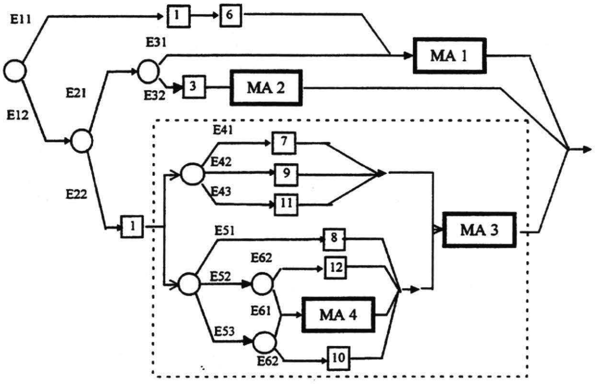


Figure 2 : Decision graph (see Figure 1 for event and action symbols and the following table for macro-actions)

Macroactions		Actions lists
MA 1	Damaged train continues service	Actions 2 and 5
MA 2	Damaged train stops service	Actions 7, 4 and 5
MA 3	Make a convoy with damaged train and next train	Actions 13, 14, 15, 5 and 16
MA 4	Empty next train at a station	Actions 17 and 8

Table 2: Macroactions and their equivalent in terms of actions

A decision graph avoids the combinatorial explosion often unavoidable in decision tree to represent all the potential paths in a reasoning when numerous choices exist. This is particularly interesting when a contextual element intervenes on few nodes of the decision tree. Comparatively, the decision graph is smaller and of the same order than the decision tree only when all the contextual elements act on all the nodes, which is not the case in our application.

In Figure 2 one can see the slashed square around the temporal branching. This represents some thing that can be see as an action ("help the damaged train") at a upper level of granularity. This simplifies the graph and permit to propose to the operators an adaptive interface : they can chose the level by developing or expanding a decision graph explaining an action, as in Sowa's (1984) conceptual graphs.

We already have studied the solving of different incidents. As a result, macro-actions give a simple picture of decision graphs and permit to point out similarities and possible shared parts between the solvings of incidents quite different.

Operators and managers at RATP have well accepted our modelling, mainly because of the clarity and understandability of the representation: it is easier to understand a set of decision graphs than the corresponding decision tree, when both of them have the same mode of investigation.

Different objects are coded to permit the handling of macro-actions and the first results are very interesting (we code in C++). The methods actually implemented concern the basic operations (creation, change, visualisation, and file management) and establishment of links among objects. Several procedures are also coded. Operators will have the possibility to modify the structure of a macro-action (i.e., a procedure) to model their own strategy for solving an incident. However, each procedure and the associated strategies will have the same structure, and there will have as many items as strategies used by operators. Concretely, a strategy is a copy of a procedure applied for an incident in a given context.

CONCLUSION

Our model has been applied to procedures of management of the four main incidential events encountered on the Parisian subway. We now look for a hybrid system associating case-based reasoning on context and decision graphs. In this paper we point out that the solving of complex incidents, which was previously represented by a complex decision tree for modelling reasoning, can find a smart representation as decision graph. Moreover, decision graphs permit to make contextual elements in the decision process explicit. This type of representation seems to us very important in domains where the decision making depends heavily on contextual factors that may intervene only on a small part of the decision process.

Acknowledgements

The SART project enters a convention between two universities, the University Paris 6 (France) and the Federal University of Rio de Janeiro (FURJ, Brazil), a contract between the University Paris 6 the company of the subway in Paris (RATP), and another contract between the FURJ and Metrô that manages the subway in Rio de Janeiro. Grants are provided by COFECUB and RATP in France and CAPES in Brazil. We also thank J.-M. Sieur at RATP, C. Gentile and M. Secron, Ph.D. students and A. Trinty and S. Roux, MsB students working in the SART project.

REFERENCES

Anderson, J. R. (1993). Rules of the Mind. Hillsdale, NJ: Erlbaum.

Brézillon, P. (1999) Context in problem solving: A survey. *The Knowledge Engineering Review*. (to appear).

Brézillon, P. & Pomerol, J.-Ch. (1998) Using contextual information in decision making. In: Chapman & Hall. Widmeyer G., Berkeley D., Brezillon P. & Rajkovic V. Eds.: *Context-Sensitive Decision Support Systems*, 158-173.

Pomerol, J.-Ch. & Brézillon, P. (1999) Dynamics between contextual knowledge and proceduralized context. *Proceedings of the 2nd International and Interdisciplinary Conference on Modeling and Using Context*, Trento, Italy, September. To be published in Series Lecture Notes of Springer Verlag.

Brézillon, P., Gentile, C., Saker, I. & Secron, M. (1997) SART: A system for supporting operators with contextual knowledge. *First International and Interdisciplinary Conference on Modeling and Using Context* (CONTEXT-97), Rio de Janeiro, Brazil, Federal University of Rio de Janeiro (Ed.), 209-222.

Brézillon, P., Pomerol, J.-Ch. & Saker, I. (1998) Contextual and contextualised knowledge: An application in subway control. Special Issue on Using Context in Applications. *International Journal on Human-Computer Studies*, 48(3), 357-373.

Carenini, G. & Moore, J.D. (1993). Generating explanations in context. International Workshop on Intelligent User Interfaces, Orlando, Florida.

Pasquier, L., Brézillon, P. & Pomerol, J.-Ch. (1999) Context and decision graphs in incident management on a subway line. *Proceedings of the 2nd International and Interdisciplinary Conference on Modeling and Using Context*, Trento, Italy, September. To be published in Series Lecture Notes of Springer Verlag.

Pomerol, J.-Ch. (1997) Artificial Intelligence and Human Decision Making. *European Journal of Operational Research, 99*, 3-25.

Sowa, J.F. (1984) Conceptual Structures: Information Processing in Mind and Machine. Addison Wesley Publishing Company.

Turner, R.M. (1999). Context-mediated behavior: An approach to explicitly representing contexts and contextual knowledge for AI Applications. *Working Notes of the AAAI-99 Workshop on Modeling and Using Context in AI Applications*, Orlando, July.

Tversky, A. (1977) . Features of similarity, *Psychological Review, 84-4,* 327--352.

RÉSUMÉ

LE RAISONNEMENT A BASE DE CAS ET LES GRAPHES DE DECISION,

APPLICATION A LA GESTION D'INCIDENTS SUR UNE LIGNE DE METRO

La gestion des lignes de métro est une tâche difficile et le contexte influe fortement sur la manière dont les incidents sont résolus. A chaque étape de la résolution d'un incident, des choix sont effectués en considérant des connaissances contextuelles. Ainsi le contexte se présente sous deux formes : le contexte procéduralisé, qui est directement pris en compte à l'étape courante de résolution, et les connaissances contextuelles. Les procédures officielles et des pratiques réelles de résolution d'incidents sont habituellement modélisées par des arbres de décision, mais ceux-ci ne permettent pas de prendre en compte l'état dynamique du contexte. Il est donc nécessaire d'étendre le modèle d'arbre de décision. C'est ce que nous proposons avec les graphes de décision, qui, de plus, prennent en compte des notions temporelles telles que la précédence et l'indépendance d'actions.

MAKING SENSE OF THE ABSTRACTION HIERARCHY

Morten Lind
Department of Automation
Technical University of Denmark
DK-2800 Lyngby Denmark
Email: ml@iau.dtu.dk

ABSTRACT

The paper discuss the abstraction hierarchy proposed by Rasmussen(1986) for design of human machine interfaces for supervisory control. The abstraction hierarchy represents the domain of human work by multiple levels of means-end and part-whole abstractions. It is argued in the paper that the abstraction hierarchy suffers from both methodological and conceptual problems. The problems are illustrated by concrete examples from the power plant domain. It is concluded that the semantics of the means-end levels and their relations should be improved by making more distinctions. Furthermore, the commitment to a fixed number of levels of means-end abstractions should be abandoned and more attention given to the problem of level identification in the modeling building process.

Keywords: Cognitive Engineering, Human-Machine Interfaces, Abstraction Hierarchies, Modeling

INTRODUCTION

The use of means-end and part-whole abstractions in plant representations for human supervisory control has been the subject of both research and application for more than a decade. Rasmussen (1986) promoted the idea in the form of the abstraction hierarchy (AH). The AH is part of a cognitive engineering approach to human-machine systems design and has been adopted mainly by the nuclear power industry, presumably because of its appeal to system engineers designing display systems for plant supervision. Furthermore, the AH is supported by empirical studies of operators fault finding strategies, and offer therefore, together with the overarching cognitive engineering framework, a promising basis for design of supervisory displays. The AH was used by American nuclear power industries after the Three Mile Island incident in their efforts to improve the reliability of the human-machine interaction and was later adopted by the Japanese nuclear industry in the conceptual development of a new generation of control rooms.

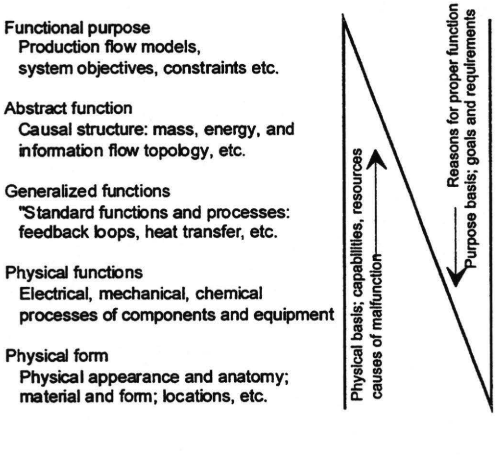


Fig.1. The abstraction hierarchy (Rasmussen, 1986)

In parallel with and partly motivated by the industrial interest, there has also been efforts within academia to develop the foundations and applications of means-end and part-whole abstractions. An example of academic research is the application of the AH for design of so-called ecological interfaces by Vicente and coworkers (Bisanz & Vicente, 1994). One of the aims of this research is to use the AH for the design of human-machine interfaces. This group has also demonstrated the application of the AH through concrete examples. Another example of academic research is the development of Multilevel Flow Modelling by Lind(1994) and his research group. The main objectives of MFM research are to develop concepts and methods for modeling of complex industrial artifacts and to use the models in conceptual design of industrial automation systems, including intelligent controls and supervisory functions for the operator (Lind, 1996). At the early stages of development MFM was presented as an articulation of the AH. However, this tight coupling of MFM to the AH proved later to be a hindrance to its development. An analysis of these restrictions led to the recognizion of the cluster of AH problems discussed below. Current MFM research (Lind, 1999) is addressing these problems and MFM is therefore no longer considered to be an integrated part of or a formalisation of the AH.

The industrials and the academics seems accordingly to share the conviction that the AH or its underlying idea of combining means-end and part-whole abstractions makes sense. In spite of these beliefs, it has actually proven to be difficult to apply the AH. However, systematic analyses of these difficulties are lacking. It is therefore not yet clear whether the very idea of using means-end and part-whole abstractions for modeling complex plants is flawed or the problems are specific to the AH. In the present paper we will identify a cluster of problems that the author has found to be sources of difficulties when building plant models using the AH.

PROBLEMS WITH THE ABSTRACTION HIERARCHY

The main principle of the AH is to describe a system on several levels of abstraction. Two types of abstraction are at work simultaneously. The *means-end* abstraction as shown in figure 1 decribe how physical resources and system functions can be organised into five levels so that each level define the means for the next upper level and define ends that are accomplished using items on the level below as the means. *The part-whole* abstractions (not in Fig. 1) decompose or aggregate items on each level of means-end abstraction. The original formulation of the AH by Rasmussen(1986) was based on studies of fault finding in electronic workshops and supervisory control of power plants. The principles of the AH were also applied by Rasmussen, Pejtersen & Goodstein(1994) to other non engineering domains leading to a generalized version. The present analysis of the AH is mainly based on the early version and shows that the AH suffer from problems even when used to represent engineering systems like power plants.

A general problem with the AH is actually its immediate appeal to common sense i.e. the fact that engineers knowledgeable especially in power systems control seems to recognize the idea of the AH without great difficulty, hence the success within the nuclear domain. This success could of course be simply explained by the fact that the AH originally was developed in the context of supervisory control of power plants. However, we also believe that the AH, more or less sucessfully, reflect tacit knowledge about engineering practices in the process control domain that also has some validity outside the narrow context of power plants. The problem is however, that the AH is lacking the expressive power to make this knowledge explicit. The common sense appeal of the AH is therefore a problem. It could lead some to the conclusion that the AH is without problems and ready to use. The analysis below shows that such a conclusion is quite premature.

The cluster of specific AH problems discussed below has been divided into two major groups; methodological and conceptual problems. The problems in the two groups are partly interdependent and the division is mainly done for convenience.

Methodological Problems

A user trying to use the AH for a particular modelling problem will be met with problems of methodological nature. One problem is that no procedure or guideline exist (to the authors knowledge) that can help in the acquisition of the background knowledge required for a modelling task. Another problem is that there is no process for model building or for revising, modifying and validating a model. These deficiencies are amplified by the circumstance, that the meanings of the different levels of means-end and part-whole abstraction in the AH are only defined in terms of prototypical incomplete examples from a few selected domains. It is therefore very difficult to apply the AH to a new or sligthly different modelling problem. A related problem is the lack of convincing arguments for the number of means-end abstraction levels. This problem will be discussed in more detail after the analysis of the conceptual problems.

The monolitical nature of the AH is another source of methodological problems. Thus, it is not clear how to combine a multiple of AH instances each representing a contextual unit of purposeful activities or how to break down a complex system in such units. A paradigm example illustrating the nature of this problem is the representation of the couplings of a production process and its control system. Production processes are physical transformations of materials or energy whereas control systems are information processing systems. It is suggested that the AH can be used to model both types of system but it is not clear how to express the fact that the information processed by the control system are representations of the physical phenomena in the production system. This coupling is established through the plant instrumentation but cannot be dealt with in an obvious way with the AH. A related problem, handled by MFM, is the representation of the couplings between a system and its supporting subsystems. These problems, indicate that the AH offer no distinctions between system categories and lacks a concept of system boundary. The problems involved in the modeling of the coupling of the plant and its control system through the instrumentation are discussed by Lind(1993) in the context of MFM.

Conceptual Problems

The hetereogeneous repository of concepts used to characterize the content of the five means-end levels is a source of confusion. According to the definitions (Rasmussen et. al.,1994), the levels includes activities, things, information, money, people, properties, constraints and prority measures. These concepts clearly refer to a variety of ideas and phenomena each belonging to different contexts of analysis, and cannot be combined without more fine grained distinctions.

Here we will mainly discuss conceptual problems that are caused by the generality of the means-end and the part–whole relations. This generality makes the AH open to interpretation, a quality that may seem as an advantage. It is actually a serious weakness because the unclear semantics disguises the real nature of the modelling problem and is a hindrance to model validation. The issues discussed in the following are; the five levels of means and ends, the concept of function, the means-end relations and the concepts of whole and parts.

The Five Levels of Means and Ends

The AH make distinctions between five different levels of means and ends. The levels of physical form and physical functions quite clearly refer to the spatial extension and behavioural characteristics of physical objects that in production plants have natural interpretations like pumps and heat exchangers. However, it is unclear whether the definition of these levels leave room for other types of entities like water and heat that according to common intuition also would count as physical means in a production plant. However, flows of water (mass) and energy are both included on the level of abstract function and not on the level of physical function. Actually, the concept of heat flow is necessary in order to describe the behaviour of a heat exchanger represented on the level of physical function and water is often used as as a means for storing energy. Of the same reason, it is not obvious how the AH handle modeling problems where a substance is both an agent and an object for action at the same time. Take for example a cooling loop. The cooling pump (agent) is transporting the water (object) and the water (agent) is transporting the energy(object). This example has three means-end levels but they do not fit naturally into the AH. The AH cannot handle this problem because substances seems to be on the level of abstract function and because it has no distinctions between agents and objects.

The inclusion of actions on the level of physical function in the AH (Rasmussen et. al., 1994) is also problematic. Most people would regard actions as genuine means (consider e.g. the following sentence "the turning of the valve by 30 degrees is a means to increase the flow of water") but actions does not to fit naturally in the same category as material objects like pumps and valves. Actions are events that take place in space and include objects and agents with physical extension but they have themselves no physical form or configuration. The relation implied in the AH between physical function and physical form is accordingly only valid for material objects. Furthermore, the intrinsic logical (i.e. not causal) relations between an action and its attributes (agent, object, means, manner, cause, result, intention etc.) considered by theories of action (Rescher, 1966) or verb semantics (Fillmore, 1968) cannot be represented fully in the AH.

Other problems occur on the level of generalized functions. The entities mentioned on his level are not physical material objects but refer to desirable outcomes of the dynamic interactions in the system and are therefore more of a process nature (e.g. cooling and control). According to the AH these processes are implemented by means of the entities described on the level of physical function. This follows from the logic of the means end relation connecting the levels of physical and generic function. Since the existence of processes then are assumed to be dependent on the existence of physical functions that again are assumed to have a spatial extension, it is seen that physical functions are considered ontologically primary to generalized functions. This implicit commitment to a thing's ontology can create conceptual problems when modeling phenomena where it is more convenient to consider processes as primary entities i.e. to adopt a process ontology (Rescher, 1996). An illustrating example from the power plant domain is the problem of representing the burning of fuel in a boiler. The combustion of fuel is a process that interacts with the air gas mixture flowing through the boiler. But this process can also be described as a flame i.e. as an entity with a physical location and form. It should accordingly be represented in the AH at the levels of physical form and function. However, even if the flame is a means for heating the boiler it is not meaningful to describe it only as a material object or thing because its existence is conditional on the burning process. It is accordingly necessary in some way to allow both material objects and processes to be at the bottom of the hierarchy. It is not clear how this should be done in the AH. A solution could be to extend the hierarchy at the bottom with a new level. But there are no principles that prescribe how such an extension should be done in a systematic way.

The Concept of Function

The distinctions made in the AH between different types of function (physical, generalized and abstract) are troublesome because the concept of function has several meanings. Proper treatment of this problem will not be attempted here because it involves a complex analysis of a whole cluster of related concepts (goal, objective, function and disposition and action). We will only point out that the AH seems to confuse functions that are ascribed to entities by convention and function that are grounded in a dispositional property of the entity in question. As an example of an entity that is ascribed a function by convention could be mentioned a coin whose function is to exchange value. An example of an entity whose function is based on a dispositional property could be mentioned a pump. The ability of a pump to serve its function (to move water) is only dependent on its physical conditions and independent on conventions. The AH invite to a confusion of the two meanings because money flow in (Rasmussen et. al., 1994) is mentioned as belonging to the level of abstract function.

This distinction between the ascription of function by convention and by disposition is not only important for understanding the differences between social and engineering systems. In fact both types of functional ascription are required in order to model the means-end relations in e.g. a power plant. As an example consider a coolant system. The function of the water circulation is here to provide cooling. The ascription of this function to the system explains why the system is there (its teleology) and is valid because the system is able to provide the function (it has the appropriate dispositions). However, if we want to explain why the water is recirculated it should be described as part of a context of value exchange where its function or role is to be a valuable object. This functional ascription is based on social conventions.

The Means-End Relations

In addition to the problems of the five levels of means and ends in the AH there are also problems with the meaning of the relations connecting them. Thus, it is not obvious that the relations between two adjacent levels in the AH are of the same nature. For example; is the relation between purposes and abstract functions of the same nature as the relations between physical functions and generic functions? Furthermore, these relations cannot be the same as the relation between physical form and physical function. Finally, it is not explained how an item on a given level is related to another item on the same level of abstraction. For example; how are two physical functions related and are their relations of the same type as the relations between two abstract functions? All these questions can only be answered by a refinement of the semantics of the means-end relation and by understanding the nature of the levels themselves.

The vague semantics of the means-end relations can in some cases lead to descriptions that are cyclic and therefore contradicting the notion of a hierarchy which is an acyclic structure. Consider for example the circulation pump in a cooling loop. The pumping (physical function) is a means of circulating the water (generalized function) that again is means of transporting energy from a heat source to a sink (abstract function). However, the pump can only circulate the water if it is not boiling and the water temperature depends on the heat transfer rate. This means that the energy transport is a means of ensuring proper conditions for the pumping. The mapping of this chain of means-end arguments into the AH requires a link between the level of abstract function (the means) and the level of physical function (the end) that seems to contradict the directions of causality (and intentionality) in the AH. The problem with this example can be resolved by clarifying the semantics of the means-end relation through a distinction between sufficient (causation) and necessary (enabling) conditionings between the means and the end. The AH does not make that distinction and therefore cannot make sense of the example. Even more problematic is the modeling of a coolant system with natural circulation where the movement of the water is caused by temperature gradients between the heat source and the sink. In this case it is questionable whether it is useful to talk about levels at all!

The Concepts of Wholes and Parts

Another cluster of problems concerns the interpretation of the levels of wholes and parts. The scope of these problems will not be explored here but only illustrated by referring to the work of Nagel (1961). Nagel analyzed different meanings of these concepts and came up with the following eight possible interpretations. A whole can be 1) something with a spatial extension and anything is then called a part of such a whole that is spatially included in it, 2) a temporal period, whose parts are temporal intervals in it, 3) any class, set or aggregate of elements, and part may then designate either any proper subclass of the initial set or any element in the set, 4) a property of an object or process, and part to some analogous property that stands to the first in certain specified relations; 5) a pattern of relations between certain specified kinds of objects or events, 6) a process, one of its parts being another process that is some discriminated phase of the more inclusive one, 7) any concrete object, and part to any of its properties, 8) any system whose spatial parts stand to each other in various relations of dynamic dependence.

It is quite clear that this variety of possible interpretations of the part-whole relation is a potential source of confusion when the actual meaning to be used in the AH is not well specified. Parts and wholes should obviously have different meanings on the five levels of means-end abstraction. On the levels of physical form and function the relation between wholes and parts is spatial inclusion (i.e. the first option in Nagel's analysis). However, on the levels of generalized and abstract functions the temporal meaning (i.e. the second option) seems more appropriate. Further confusion is introduced because one of the levels of part-whole abstraction in the AH is labeled the level of "function units" (Rasmussen, 1986, p. 119). This labeling confuses concepts of function belonging to the means-end distinction with concepts belonging to the part-whole decomposition.

DISCUSSION AND SYNTHESIS

The analysis above has shown that the AH suffer from both methodological and conceptual problems. The conceptual problems have been demonstrated through detailed examples from the power plant domain. In the following we will discusss the means-end concepts and specific problems of the AH in the broader contexts of decision making and modeling methodology. The following two questions will be addressed: 1) What are the advantages and uses of means-end concepts in modeling complex systems? and 2) How are levels of means and ends identified? The first question is adressed by Rasmussen but not completely answered and the second has not been raised. The discussions reveal that the AH in addition to the specific problems mentined bove also suffer from more basic problems both when seen in an application perspective and as a modeling framework.

Advantages and uses of Means -End Concepts

Means-ends concepts play an essential role in theories of action and practical reasoning (VonWright, 1963) and has potential application in many domains. Rasmussen(1984) argue that the means-end representation instantiated in the AH provides a systematic framework for identifying and evaluating alternative courses of action and that the AH in this way can reduce the complexity of decision making in supervision and control of anomalous plant situations. The main feature of means-end representations (an thereby also the AH) is that they define opportunities for decision making and support an organized sequential decision process.

Rasmussen combines the AH with a model of decision making. This decision model provides in principle a decomposition of the supervisory control problem into three basic decision problems in supervisory control each coping with a separate aspect of uncertainty. The three problems are state identification, goal selection and action planning. In state identification the problem is to select between alternative interpretations of the system state caused by multiple, insufficient or uncertain observations. In goal selection the problem is to choose or make compromises between multiple contradicting goals. The problem in action planning is to select among alternative courses of action. The three problems cope with three distinct resources (and preferences) of the decisionmaker – the set of possible observations and perceptions of the system state, the range of possible goals and the possibilities for action. The AH is proposed as a framework for solving the overall supervisory control problem, but it is not clear whether the AH can support all three subordinate decision problems. The AH seems to be most fit for the action planning problem since the entities described on the level of physical function and functional purpose represent means and ends of intervention. However, other representations seems to be required for state identification where the problem is to manage multiple and possibly conflicting means of observation and state interpretation. Rasmussen emphasizes the interdependence of the three decision problems in supervisory control and suggest that the causal and intentional reasoning processes supported by the AH provide a framework for their integration. However, even though the decision problems are interdependent they may require application of three separate means-end representations because the problem involves the management of different categories of means and ends.

Identification of Levels of Means and Ends

It is clear from the discussion above that the quality of decision making is directly dependent on the problem framing provided by the levels of means and ends i.e. the number and semantics of the levels and the means-ends relations. Improperly defined levels will lead to inefficient or directly wrong decisions. If the levels of description are too abstract decision alternatives may be overlooked. If the levels are too detailed they may include irrelevant alternatives and may lead to uncertain decisions. It is therefore essential that the methodology used to build the means-end representation has explicit principles, rules or guidelines for identification of levels and their semantics. Actually changes of representation of an artifact through redefinition of levels of descriptions may sometimes be a necessary step in a problem solving process. The problem framing provided by the levels of description may be a hindrance rather than a help in situations that call for the creation of new opportunities for decision making.

The AH framework does not present the number of levels and their semantics as subjects for decision in the modelling proces. The AH with the five levels of means-end abstraction seems to pretend that there is no problem of levels at all. A user will therefore tend to regard the number of levels of means-end abstraction in the AH as paradigmatic and may find it difficult to fit the levels to a given problem. The levels should rather be seen as resulting from a process of framing that may have different outcomes in different modeling situations. In other words, the number and contents of the levels and their relations cannot be defined in general but are features that characterize a particular modeling domain or even a specific problem. The problem is to understand how levels are constructed i.e. to know the criteria used in the selection of perspectives and particular aspects of a system or a situation and how these perspectives and aspects are combined into a representation that can be used to solve a given decision problem. Only through such an understanding can means-end concepts be used to construct models that makes sense.

ACKNOWLEDGEMENT

The present work has been supported by the Danish Foundation for Basic Research through a contract with the CHMI Center for Human-Machine Interaction.

REFERENCES

Bisantz, A. M. & Vicente, K. J. (1994). Making the Abstraction Hierarchy Concrete. *Int. J. Human-Computer Studies*, 40, 83-117.

Fillmore, C. (1968). The Case for Case. In: E. Bach and R. T. Harms (Eds), *Universals in Linguistic Theory.* New York: Holt, Rinehart and Winston.

Lind, M. (1993). *Functional Architectures for Systems Management and Control.* In Lind, M et. al., Interactive Planning for Integrated Supervision and Control in Complex Plant. Final report from CEC JRC project: 4937-92-08-ED ISP DK.

Lind, M. (1994). Modelling Goals and Functions of Complex Industrial Plant. *Journal of Applied Artificial Intelligence*, 8, 259-283.

Lind, M. (1996). Status and Challenges of Intelligent Plant Control. *Annual Review of Control*, 20, 22-31.

Lind, M. (1999). *Plant Modeling for Human Supervisory Control.* Submitted for publication.

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation.* New York: Hartcourt, Brace & World, Inc.

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction.* New-York: North Holland.

Rasmussen, J., Pejtersen, A. M. & Goodstein, L. P. (1994). *Cognitive Systems Engineering.* New York: John Wiley.

Rescher, N. (1966). Aspects of Action. In N. Rescher (Ed), *The Logic of Decision and Action.* Pittsburgh: University of Pittsburgh Press.

Rescher, N. (1996). *Process Metaphysics – An Introduction to Process Philosophy.* New-York: State University of New York Press.

VonWright, G. H. (1963). *Norm and Action.* London: Routledge & Kegan Paul.

RÉSUMÉ

DONNER DU SENS À LA HIÉRARCHIE D'ABSTRACTION

Cette communication porte sur la hiérarchie d'abstraction proposée par Rasmussen (1986) pour la conception d'interfaces homme-machine dans le contrôle-supervision. Cette hiérarchie représente le domaine du travail humain selon des abstractions de niveaux multiples en termes, soit de fins et moyens, soit de tout et parties. On montre que la hiérarchie d'abstraction pose des problèmes méthodologiques et conceptuels. Ces problèmes sont illustrés par des exemples concrets empruntés aux centrales électriques. On conclut que la sémantique des niveaux de fins et moyens, ainsi que de leurs relations, nécessite des distinctions supplémentaires. De plus, le nombre fixe de niveaux de cette hiérarchie doit être abandonné et davantage d'attention doit être accordée au problème de l'identification des niveaux dans le processus de modélisation.

MOTS CLÉS : Génie cognitif, Interfaces homme-machine, Hiérarchies d'abstraction, Modélisation.

STUDY ON OPERATOR'S KNOWLEDGE MODEL BY SPEECH ACT ANALYSIS

Kazuo Furuta* and Shinji Yoshikawa**
* Institute of Environmental Studies, The University of Tokyo
E-mail: furuta@k.u-tokyo.ac.jp
** Japan Nuclear Cycle Development Institute
E-mail: syoshi@t-hq.jnc.go.jp

ABSTRACT

This work is to propose an operator's knowledge model based on observation of human performance in process control. An experiment was carried out, and subjects' cognitive processes were described as logical formulae so that knowledge used were elicited explicitly. The analysis revealed that there are primarily four types of knowledge: state, goal, causality, and configuration. Subjects' cognitive processes in different experimental scenarios were characterized by classifying their utterances into the knowledge types and then temporally tracing the processes. Consequently a knowledge model that consists of four spaces for these knowledge types and their interrelations will be proposed.

Keywords

Plant operator, human modeling, knowledge model, protocol analysis, speech act theory.

INTRODUCTION

The more modern technological systems get large and complex, the more difficult it becomes for humans to understand their mechanism and behavior. Various support systems for diagnosis, operation, maintenance, or training have been studied and developed in process control to mitigate this problem. Such support systems should, however, not only respond correctly but also communicate with humans so that they will not make new black boxes. A human model, which is a model of human cognitive behavior, is a key basis for developing support systems that can communicate with humans.

A human model comprises three components: process model, knowledge model, and control model. This work deals with the knowledge model, which describes the substance and the representation of knowledge used in human cognition, in particular what kinds of concepts, relations, and vocabularies are used while thinking and acting. The knowledge model is a similar concept as ontologies discussed in knowledge engineering.

We have already proposed a knowledge model of a plant operator that consists of two layers of Task Hierarchy (TH) and Qualitative Causality (QC) for development of a training support system (Furuhama, Furuta, & Kondo, 1995). In this model, knowledge on operator's tasks is described by goal-mean relations in the TH layer, while knowledge on system behavior is described by qualitative cause-consequence relations in the QC layer. Though this model have been validated partly by experiments, it has some deficiencies as a generic model, such that it does not consider the interpretation stage in human cognitive processes.

From the above backgrounds, the aim of this work is to propose a generic model of operator's knowledge based on empirical evidence of human performance. In this paper we will explain the experiment and protocol analysis we conducted for eliciting knowledge used in plant operation, and then propose a framework of knowledge based on substantial classification of knowledge extracted by the analysis.

EXPERIMENT

We analyzed the protocol data of an experiment that one of the authors had already reported (Yoshikawa, Ozawa, Koyagoshi, & Oodo, 1996). The experiment was carried out using the full-scope training simulator of the prototype fast breeder reactor, Monju. The subjects were two crews of three operators, and the instructor operating the simulator in the instructor room served as a field operator. Subjects' behavior was recorded using four remote-controlled cameras, eight wireless microphones, two video recorders, and four mini-disc audio recorders. Subjects' operational actions on the plant were recorded by the logging function of the simulator.

Two scenarios were assumed in the experiment. The first one is double leaks in a primary sodium coolant loop, which is due to simultaneous ruptures of a pipe around the exit of a primary circulation pump and a tube of the Intermediate Heat Exchanger (IHX) in the same loop. Some of the symptoms to be observed in the first rupture are shadowed by those due to the second rupture, because the sodium leak rate of the second rupture exceeds that of the first. The probability of this scenario is extremely low, and it is included neither in the design basis events nor in the training scenarios. The second experimental scenario is loss of steam supply to a feed water heater, which is caused by abrupt closure of the check valve in the steam line from a high pressure turbine to the heater. In this scenario evolution of anomaly is moderate, and no alarms annunciate for several minutes after event initiation. The symptoms to be observed are oblique for the operators to identify the root cause of the anomaly so that the operators need a lot of mental simulation for event identification. The experimental runs lasted about 40 minutes for the both scenarios till the subjects successfully identified the events and started counter actions.

PROTOCOL ANALYSIS

Subjects' cognitive processes were reconstructed from their verbal protocols considering context of statements and actions to elicit what knowledge were used. First we tried to analyze the protocols by expert judgement based on predefined criteria. Objective and reliable analysis was however hard by such a naive method; we adopted a stronger method to describe subjects' cognitive processes that formed the observed actions and utterances by formal representation. Analysis by this method is almost the same as hand simulation of subjects' behavior, where execution of each cognitive step is controlled by the analyst following the experimental records. It is expected that the formalities introduced by the method contribute to reducing logical ambiguities in description and improving objectivity of analysis, while expressiveness of the formal representation is limited and the analysis can be still subjective.

In this analysis a method is required to represent cognitive state of an agent separately from the real state of the world. Having introduced a modal operator BEL, "A BEL P" implies that an agent A believes a proposition P. The operator BEL can be used recursively to represent one's belief on another's belief like "A BEL B BEL P", which means that A believes that B believes P. Another operator WANT is used to represent one's goal or plan. That A wants P as a goal to be achieved or as a plan to be executed is represented in combination with BEL like "A BEL A WANT P". Belief on the possibility of a proposition is represented by another operator $\Diamond$ like "A BEL $\Diamond$P", which means that A believes that P is possibly true. Propositions not preceded by $\Diamond$ is thought necessarily true. Some axioms hold for these operators such as "A BEL A BEL P $\leftarrow$ A BEL P", but they are not important here.

Since the subjects are well-trained experts of plant operation, it is assumed that they behave rationally following deliberated plans. According to the formal theory of planning, human action is the result of application of an operator, which is defined by effects, selection conditions, preconditions, and body. In planning of action, an operator is chosen whose effects include the specified goal and whose selection conditions are satisfied, and it is then added to the plan. If some of the preconditions are not satisfied, achievement of the unsatisfied preconditions become sub-goals of the plan. In addition to planned actions, context-driven actions, which are triggered when the preconditions match the present situation, were took into account in protocol analysis.

We thought speech as a kind of action that an agent performs to try to change hearer's beliefs following the speech act theory (Cohen & Perrault, 1979). We can thereby describe and analyze plans for both operational actions and speech acts on the same framework. Plans related to domain specific knowledge of plant operation were distinguished from plans related to common sense knowledge of discourse. The model of domain specific knowledge is the subject of this study, but analyzing the former is necessary also to reveal the latter. Typical operators for the latter are the Inform speech act and the Request speech act. If an instance of these speech acts is found in subjects' protocols, speaker's goal can be inferred by inversely applying speech act operators. The definition of speech act operators used in this study are almost from Allen and Perrault (1980). The same analysis was performed also on the observed actions to infer subjects' intended goals.

EXPERIMENTAL RESULT

The knowledge elicited by reconstruction of subjects' cognitive processes could be categorized by key concepts and fundamental relations they imply. The following four spaces of a knowledge model have been defined for representing knowledge types observed frequently.

State Space

Identification of the system state from observed pattern of trends in system parameters, symptoms, is a typical way of thinking. Figure 1 shows the logic that derived the subject C's utterances in the excerpt:

C: Sodium leakage! (alarm)
B: Yes, ... I will check it.
C: Hey, level change and, ..., Mr. A, will you check the atmosphere temperature?

In the above, association between the state *leak(Loop1A)* and the set of symptoms {*alarm(leak(Loop1A)) is on, level is low, atm_temp is high, ...*} is used. Since the subject C spoke immediately after having heard the report on the annunciating alarm and since the event *leak(Loop1A)* is a well-trained scenario, it seems the association had not been derived but preexisted as knowledge. Such knowledge is used for state identification in a way, "◇State_A ← Symptom_X", while the same knowledge is used for prediction of unobserved symptoms in the opposite way, "◇Symptom_X ← ◇State_A". We will call the region where knowledge of this type are to be represented the state space.

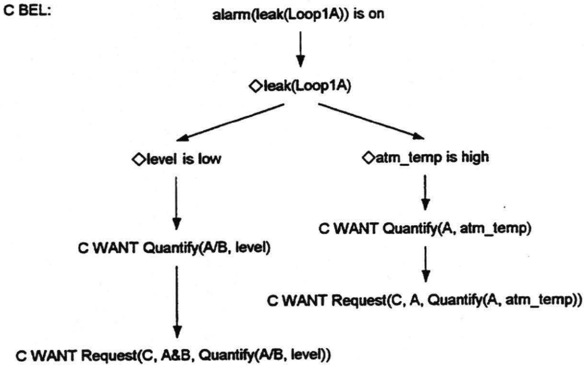


Figure 1: An example of logical formulation of speech acts.

The states referred to by the subjects include not only root causes of anomalies, but also more general or abstract states of the plant that are useful for the purpose of plant operation. The subjects distinguished the normal operating condition, abnormal conditions in some subsystems, specific situations that request operators to initiate some action, and so on. These states form a hierarchy, where a general state can be classified into more specific states. The associated symptom set is the guideline for classifying and identifying a particular plant state. A specific state is associated with a more specific and larger symptom set than its general antecedent. Symptoms are usually expressed as qualitative interpretation of plant parameters, and the subjects preferred qualitative expression much more than quantitative expression. In each experimental scenario, reference to qualitative expression was found in 143/102 utterances, while quantitative expression in only 22/18 utterances.

Goal Space

The subjects used knowledge on their own tasks as well as on the target plant for deciding what they should do. The following conversation is an example.

B: Well, just SID? It's low. Shall we demount the DPD filter and carry out chemical analysis?
C: Yes.
B: Whether or not there is deposit.
C: Can you check it?
B: Call the chief, and enter the CV, and demount the DPD filter, and ....

A concrete procedure to achieve the goal of confirming suspected sodium leak was recalled or planned here. Goal-mean relations among tasks are the basis of knowledge in action planning as above, and these relations also form a goal-mean hierarchy of tasks. The goal space is the field where such knowledge on goal-mean relations between operator's tasks and system functions are to be modeled.

Causality Space

If the situation the subjects faced was unfamiliar, they used more fundamental knowledge on behavior of the system for problem solving. The following is an example of conversation in such a process.

A: The electric power won't increase just because the flow increased. Then, the steam extraction went wrong, didn't it?
B: Maybe. The power increased and the feed water temperature got low to the extent that the extraction flow decreased and went to the turbine. Or if this went wrong before, ...

Decrease of the extraction flow caused increase of the turbine flow, and then increase of the electric power, but at the same time it resulted in the low feed water temperature. In this inference process, knowledge of qualitative causalities between plant parameters were used, which can be described as positive correlation, "If parameter *P* increases (decreases), then parameter *Q* increases (decreases)," or as negative correlation, "If parameter *P* increases (decreases), then parameter *Q* decreases (increases)." We will define the causality space for representing knowledge on such qualitative causalities between parameters. Some physical law, like "*steam_flow = extraction_flow + turbine_flow*" for the above example, must exist behind qualitative causalities. Since physical laws themselves, however, were seldom referred to by the subjects, it seems that they were compiled into qualitative causalities before used in reasoning.

Configuration Space

The most basic knowledge the subjects used in problem solving were those on the structure and properties of the plant system. The following represents a good example of such an occasion.

B: The extraction line is here, isn't it? This line is from the auxiliary heater ...
A: Does it include any bypass line?
B: No.
A: So, is this an AV?

In the loss of extraction scenario, search following connections between plant components as above continued for a considerable length of time, because the symptoms observed were not straightforward for identifying the event and the structure around the failed component was complicated. Any expression on existence, connection, location, or attribute of a plant component is typical of this category. An attribute here means some design specification or property of a plant component such as the capacity of the Over Flow Tank (OFT), which was actually referred to frequently in the double leak scenario. Knowledge on plant configuration form a hierarchy of different levels of detail from the whole plant down to parts of components. The subjects used to think in the level sufficiently detailed for the present task. We will call the region to model such knowledge on the plant system the configuration space.

Transition Between Different Spaces

Subjects' protocols show that the subjects often performed inference over different types of knowledge and that the four knowledge spaces described above are not isolated but connected. For example, if some hypothesis on the plant state has been created, a particular action may be recalled and triggered. The state *leak(Loop1A)* in Figure 1 led to a task of *Check(F, deposit(Na, DPD_filter))* via prediction of *deposit(Na, DPD_filter)*. It is transition of inference from the state space to the goal space. In another example, as a result of the inference exemplified in the subsection on the causality space, a symptom set *{electric_power is high, feedwater_temp is low}* was derived for identifying *extraction_flow is low*. Such derivation of teleological knowledge from etiological knowledge is called knowledge compilation in artificial intelligence community. Inter-links between different knowledge spaces enable humans to carry out online knowledge compilation in problem solving. Similar transitions between different spaces were observed from the causality to the goal, from the goal to the causality, and from the configuration to the causality space.

Temporal Trace of Cognitive Processes

Global flow of subjects' cognitive processes can be figured out by tracing temporally the type of knowledge used during the experimental runs as shown in Figure 2 for the double leak scenario. A small dot on a horizontal line represents an utterance related to each knowledge space, and a vertical arrow represents transition of inference between different knowledge spaces. This diagram shows subjects' cognitive processes featured by five phases as follows.

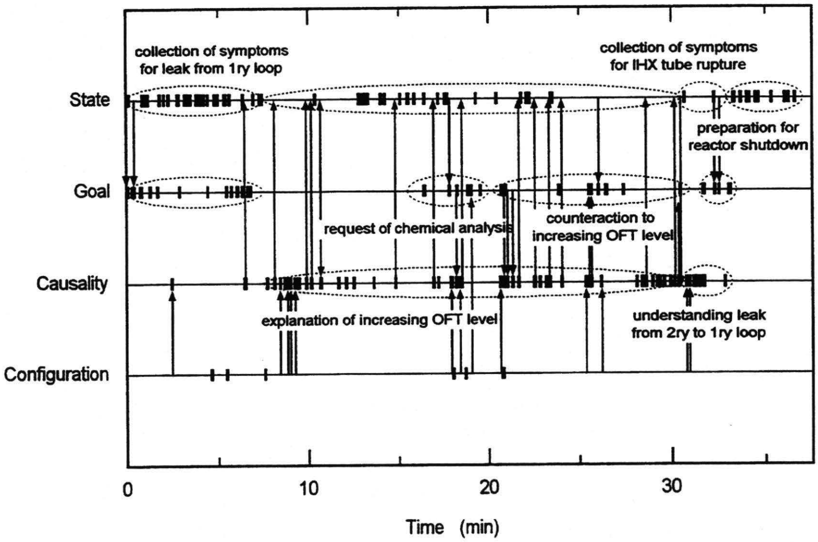


Figure 2: Temporal trace of cognitive processes in the double leak scenario.

1. The subjects first tried to confirm the event the alarm implied by observing symptoms expected in case of sodium leak from a primary loop. In this phase the subjects worked primarily in the state space.
2. Since increase of OFT level was not understandable by state-symptom association, they tried to explain this curious plant behavior using knowledge of the causality space, but they failed.
3. At least the subjects called a field operator and asked him to confirm leak from a primary loop by chemical analysis. This task was planned in the goal space.
4. Then the subjects were obliged to consider some counteractions to the increasing OFT level in the goal space.
5. Finally the subjects found symptoms indicating IHX tube rupture, understood the plant behavior using knowledge of the causality space, and decided to shutdown the reactor.

In the loss of extraction scenario, reasoning in the causality space began just after detection of anomaly, because there were no significant symptoms for event identification. Search in the configuration space appeared frequently in a time interval from 10 to 24 minutes. Comparison with the double leak scenario reveals that the two scenarios imposed somewhat different cognitive demands on the subjects.

CONCLUSION

We performed an experiment and speech act analysis of plant operation to clarify the substance of operators' knowledge on the plant system and their tasks and to propose a generic model of knowledge. We logically formulated subjects' cognitive processes based on the theories of planned action and speech act. The analysis showed that the substance of knowledge used in the experiment could be classified into four categories and their interrelations so that we have proposed a knowledge model that comprises four spaces of knowledge: state, goal, causality, and configuration. Temporal trace of subjects' utterances from a viewpoint of this model could characterize their cognitive processes for the experimental scenarios.

The substance of knowledge used were classified and represented mostly by the four knowledge spaces, but there were some knowledge that were out of the scope of this taxonomy. One of them was knowledge on quantitative relations between parameters. Subjects' utterances related to quantitative inference were very rare and restricted so that the substance of quantitative knowledge was not clear. Another type not covered by the model was knowledge on cooperation of team performance, which was implicitly assumed in the most speech acts but not explicitly stated. Some extension of the model will be necessary for these knowledge types in the future study.

REFERENCES

Allen, J. & Perrault, C.R. (1980). Analyzing Intention in Utterances, *Artificial Intelligence, 15*, 143-178.

Cohen, P.R. & Perrault, C.R. (1979). Elements of a Plan-Based Theory of Speech Acts, *Cognitive Science, 3*, 177-212.

Furuhama, Y., Furuta, K., & Kondo, S. (1995). *A Methodology to Represent Plant Operator's Mental Model for Training Support System*, Paper presented at the *5th European Conf. on Cognitive Science Approaches to Process Control (CSAPC '95)*. Espoo, Finland, Aug.

Yoshikawa, S., Ozawa, K., Koyagoshi, N., & Oodo, T. (1996). Knowledge Dependency Analysis of Plant Operators Using Consistent Protocol Formulation. In P.C. Cacciabue & I.A. Papazoglou (Eds.), *Probabilistic Safety Assessment and Management* (pp. 21-26). London, UK: Springer (4 authors and more).

RÉSUMÉ

L'ÉTUDE DU MODÈLE DES CONNAISSANCES DE L'OPÉRATEUR PAR L'ANALYSE DES ACTES DE LANGAGE

Ce travail propose un modèle des connaissances de l'opérateur fondé sur l'observation de l'activité humaine dans le contrôle de processus. Une expérience a été conduite et les processus cognitifs des sujets ont été décrits sous la forme de formules logiques de sorte que les connaissances utilisées soient explicitement extraites. L'analyse révèle qu'il y a principalement quatre types de connaissances, sur les états, les buts, la causalité et la configuration. Les processus cognitifs des sujets dans différents scénarios expérimentaux ont été caractérisés en classant leurs verbalisations selon ces catégories et en représentant leur décours temporel. En conséquence, un modèle de connaissances est proposé qui inclut quatre espaces pour les quatre types de connaissances, ainsi que leurs interrelations.

MOTS CLÉS : Opérateur de contrôle de processus, Modélisation humaine, Modèle de connaissances, Analyse de protocoles, Théorie des Actes de Langage

INTRA-NATIONAL CULTURAL DIFFERENCES RESULTING FROM HUMAN FACTORS ANALYSIS OF COMPLEX SOCIO-TECHNICAL SYSTEMS

P. C. Cacciabue, P. Ponzio, M. Bacchi

European Commission,
Joint Research Centre,
Institute for Systems, Informatics and Safety
21020 Ispra (Va), Italy
pietro.cacciabue@jrc.it - paola.ponzio@jrc.it - monica.bacchi@jrc.it

ABSTRACT

This paper presents a combined theoretical-field study of a large company with the aim of evaluating important factors affecting the behaviour of operators and devising a number of recommendations for improving the safety standards. Moreover, as the study covered a very broad spectrum of factors and issues, an additional finding, not initially planned, resulted with respect to the different safety culture showed by operators in different regions of the same nation. This result turned out to be particularly significant and could offer further insight for the development of new policies within the organisation.

Keywords: Human Factors data, accident investigation, field studies, socio-technical systems, organisational studies.

INTRODUCTION

The consideration of the interaction of human-machine-socio-technical context is emerging as the critical issue to be studied and developed, for reaching the critical goals of "modern" technology of improving *production*, *efficiency*, and *effectiveness*, while maintaining high safety standards and environmental friendly considerations. A clear example of this new type of demand is shown by the way in which automation design and application is evolving. Indeed, *automation*, which is the most relevant measure introduced in the last three decades by development of technology, is moving towards the application of "human centred design" concepts that are beginning to be applied by manufacturers and users in designing control systems and interfaces (Billings, 1997).

Even if a leading position in studying this issue has been taken by the aviation industry, (Edwards, 1972; Wiener & Nagel, 1988), these concepts and goals are shared by many different domains such as land transport, energy production, manufacturing, and process engineering. However, the rate at which technology has developed differs and depends on economical and political reasons. It is thus natural to attempt at "transferring technology" concepts and implementations between different application domains. The process of *technology transfer* is not only depending on the compatibility between systems and processes, but implies much wider socio-technical considerations (Sheridan, 1992).

The analysis of Human Factors (HF) and social contexts associated to a specific field is thus central both for automation development and for technology transfer. In other words, either for generating a valid design of an automatic system or for applying in a certain domain the technology principles developed somewhere else, a solid analysis of the company/organisation/working environment/task (human factors study) is necessary.

The transport domain is a stereotype example of such field where automation is playing a crucial role for improving control and efficiency. However, human factors concepts are differently developed depending on the type of transport. In the aviation domain, the human factors and socio-technical considerations are widely applied and fully integrated in the processes of design, training, safety assessment, and accident investigations. In other transport systems, e. g., railways, maritime, and road, these studies are not yet integrated by manufacturers and users and are being applied only for audit and assessment purposes.

In this paper we will present an approach chosen for analysing a large national railway organisation with considerable technology and social problems. We have performed the assessment of front line operators (train drivers) to identify reasons of an apparently decrease level of performance, that could have led to recent accidents within the company. We will discuss the models and techniques selected for the analysis, exploiting concepts and models applied in the aviation and nuclear energy in the light of technology transfer principles.

The results obtained from the analysis applied to a relevant population of subjects, distributed over a vast longitudinal geographical area, will be presented by discussing the recommendations that could be drawn from the study and by highlighting specific differences noticed at level of sub-culture exhibited by train drivers.

METHODOLOGY APPLIED FOR STUDYING THE ORGANISATION

The study of the Organisation and of the Human Factors has followed a "classical" approach based on the contribution of three major lines of analysis:

1. The theoretical and methodological approach; coupled with
2. A set of field analyses, based on interviews and workshops; and with
3. The study of written guides and procedures and of accident reports statistics and retrospective studies.

These three lines of the analysis have not been carried out in sequence, but in a progressive iterative manner, with frequent feedback and reviews. The leading theoretical standpoints were continuously compared with real working conditions and facts and with current rules and regulations governing the everyday life of the organisation. In this way, the theories that have been applied could be "validated" and/or "adapted" to the actual contexts of the specific organisation under study.

Figure 1 shows in a graphical form the combination of these three lines of development and their major interconnections.

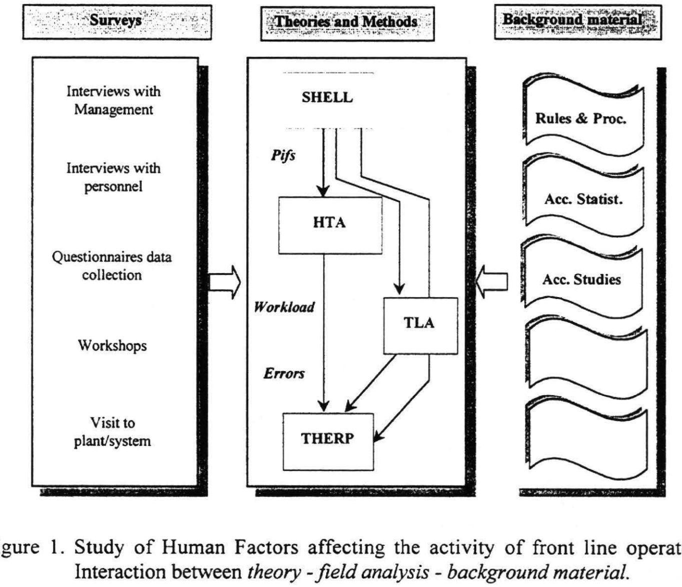


Figure 1. Study of Human Factors affecting the activity of front line operators. Interaction between *theory - field analysis - background material.*

The analysis phase started by performing some preliminary interviews with a number of operators, as well as with the organisation managers in order to evaluate a *reference model* of human-machine interaction and to know as much as possible about the reality of the organisation.

This process demanded to select a reference model for the analysis of operators and their working context. The choice fell on the so-called SHELL (Edwards, 1972; Hawkins, 1975), because of three major considerations:

1. SHELL is the reference model adopted in domains strongly affected by human factor issues, such as the domain of aviation, for the classification of accidents and incidents;

2. SHELL is a consolidated framework of reference, developed in the early 70ies, which has been validated and widely applied in other real working context studies; and, finally,
3. SHELL was known to the organisation experts and had already been used in the past for similar studies.

The friendliness of the SHELL model and its applicability to the case study was initially tested during the first Workshop where the definitions of the four areas of the model were adjusted in a slightly different manner than the original meaning of each element of SHELL. In this way it was expected that the interaction with the front line operators would have been more efficient allowing more fruitful interviews. Indeed, the model allowed to concentrate the attention on what the operators really considered to be problematic features and to easily collect the information in order to produce practical recommendations.

The first set of interviews has mainly considered logistic aspects, problems with shift planning, procedures and rules and relationship with colleagues according to the chosen theoretical model. Another important result of the first interviews and workshop was to develop an initial set of factors affecting performance and behaviour, called *Performance Influencing Factors (PIF),* and to devise the guidelines for the interviews and a short questionnaire, by which to collect a written feedback on the most crucial aspects found out.

The second step of the analysis was the identification of the critical tasks with respect to safety, by implementing the theoretical instrument of the so-called *Hierarchical Task Analysis (HTA)*. This methodology allows to represent in a simple structured format, of a "tree like", the sequence of tasks to be carried out by the operator starting from the high level tasks and arriving at detailed specific action level. In this exercise of task identification, priority was given to the practical aspect of the performance of tasks rather than the theoretical one (Kirwan & Ainsworth, 1992; Redmil & Rajan, 1997).

In order to represent formally the time distribution of work and associated workload *Time Line Analysis (TLA)* has been chosen, by which it is possible to represent, in a simple graphical format, the sequence of actions to be carried out by the operators during the performance of a task. In this way, it is possible to compare the expected workloads. Successive interviews and workshop had the objective of a deeper evaluation of the *PIFs* table already devised during the previous activity and further modified by the *HTA* and *TLA* analyses, and allowed to start discussing the issue of error making with operators so as to obtain interesting feedback. An example of *TLA* analysis is shown in figure 2.

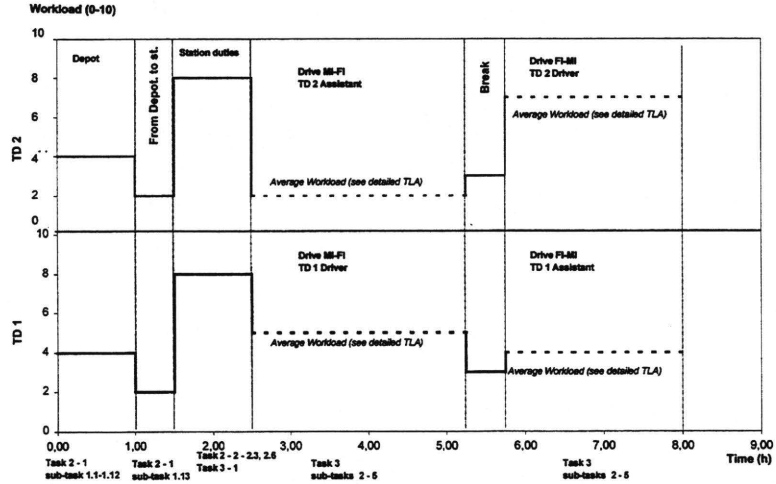


Figure 2. Time Line Analysis for a 9-hour period of a standard shift

The interviews focused therefore on the ability of the interviewees to speculate on possible accident sequences and consequences of human errors and on their likely causes. This lead to the identification of an initial set of sequences, which were structured according to the theoretical framework denominated THERP (Technique for Human Error Rate Prediction). This theory is normally used in Human Reliability Assessment studies to calculate probabilities of human errors associated with the performance of tasks. However, in this case we only applied the binary tree structure of THERP to represent combination of human errors and possible consequences in the evolution of the events. No probability analyses were carried out, as this subject was outside the scope of the study and it also represent the most critical aspect of the methodology. The graphical structures of sequences depicted by THERP were further developed with discussions with operators and expert and were then compared with the *Background Material* contained in the findings of accidents from '92 and in retrospective studies concerning 30 incidents occurred in the recent past.

RESULTS AND RECOMMENDATIONS

The recommendations developed from the study could be grouped in two main areas. The first one concerned the actual interfaces between drivers, and the working personnel of the company, with the "hardware" of their everyday work. The second area focused on the more subtle human and social aspects of the organisation.
As for the human-hardware interface, three main specific sectors calling for improvement were identified:

1. The technological interfaces, which were not considered sufficiently friendly or designed according to human-centred principles;
2. The maintenance process, which was found insufficiently structured to grant high safety standards; and
3. The physical working environment, which showed strong variation between modern and old cabins.

Detailed aspects of improvement for each of these sectors were then pinpointed and notified to the company.

As for the human and social area, the recommendations were organised according to the "4-Ps" approach (Degani & Wiener, 1994), which structures the organisational aspects at 4 levels of decisions according to the objectives and responsibilities within an organisation. In short, top decision-makers and management should be concerned with the Philosophies, and with the way in which these are implemented (Policies), while the actual performances and everyday work relate to Procedures and Practices. The rationale for synthesise the recommendations in this way is that the 4-Ps match very well with the SHELL model as it was applied for representing the drivers domain. In particular:

1 At level of Philosophies, it has been suggested to make more explicit and visible the process of communication between top management and train drivers.

2 At level of Policies, two suggestions were made:
  - a more reliable way for exploiting the communication between management and operators; and
  - a more complete process of training.

3 At level of Procedures, a different way of structuring shifts and working programmes.

4 Finally, at level of Practices, a greater involvement of the drivers in the definition of the working procedures was identified, so as to consider the actual performance of work.

As for the case of human-hardware interfaces, also in the sector of social aspects, these recommendations at high level of scope, have been complemented with specific suggestions for practical implementation.

INTRA-CULTURAL DIFFERENCES WITHIN THE SAME NATIONAL ORGANISATION

As the study concerned a remarkable population of subjects belonging to different geographical areas, an additional and significant finding, not initially planned, resulted from the analysis with respect to the different (safety) cultures showed by the operators in different regions of the same nation.
This result turned out to be very interesting and was initially due to personal perception of the interviewers. Considering the three major geographical areas where the interviews took place, the interviewers had the opportunity to verify a series of differences and peculiar characteristics of operators of the Southern, Northern and Central regions. As an example, the "special availability" shown by drivers of the South towards the observers, and the "high professional behaviour" of drivers from the North, are amongst the most relevant elements that showed, at first impact, the different cultural aspects within the same organisation.
As organisational culture is developed through the interaction between external and internal factors mutually influencing each other, the environment, the operators'charachteristics, and the way to relate to the management, should be considered important variables influencing the organisation and its safety culture.
In general, shared experiences within groups provide members with materials for sense making that others do not share.
As in this case the organisation covers an extended geographical area, specific differences exhibited by the operators at the sub-cultural level can be easily pointed out as a natural consequence of the organisation dimension, which contributed to the creation of regional sub-groups with their own distinctive cultural profile, own characteristics and common shared mental models.

The initial perception of the observers and the above theoretical considerations about safety culture were further validated by the results of the questionnaire. In particular, the questionnaire revealed some peculiarity in the answers given by operators according to their specific working areas especially as regards to the different level of agreement or disagreement expressed.
Though in general, the comparison of the frequency distribution obtained on the entire national sample with those resulting from each single area did not reveal such remarkable differences, some issues were perceived as more problematic and critical for operators belonging to an area rather than to another. This feature is evident when comparing each relative frequency distribution, with the one of the other areas considered.
Consequently, in order to highlight the regional differences and to evaluate whether these differences demand special attention for developing recommendations, we focused on the latter. Indeed, different positions expressed by operators towards some issues may be very important for further interventions as they allow to better consider possible differences coming out in different regions.

In particular, the questionnaire underlined the following main aspects:

- A different position was taken by the operators as regard to the role of the Line Instructor and his/hers personal involvement in the operators' working problems. The questionnaire showed that this role has been more easily accepted and implemented in the South and Central areas while, it has had some operative difficulties in the north part of the nation, probably because of the extensive dimension of the organisation there and the low ratio instructors vs. drivers.
- Effects on drivers' safety related performance of the changes to the technical environment - including increased train frequencies and number of tasks to be performed – were perceived as aspects able to influence the operators' workload and fatigue primarily for the south and central regions.
- The operators of the northern area expressed an agreement related to the availability of the middle management in solving possible operational problems rather than in the other two areas (Chart 1 and 2).

The following charts present the percentage of the evaluation of item 9 of the questionnaire that takes into consideration both regional differences (chart 1) and percentage related to the entire sample (chart 2).
Chart 1 reveals a quite clear difference in the opinion expressed by the northern operators (partial + total agreement ≥ 60%) in relation to the central (partial + total disagreement ≥ 50%) and southern ones (partial + total disagreement ≥ 60%) as regards to the management involvement in the everyday working problems operators may have to face.

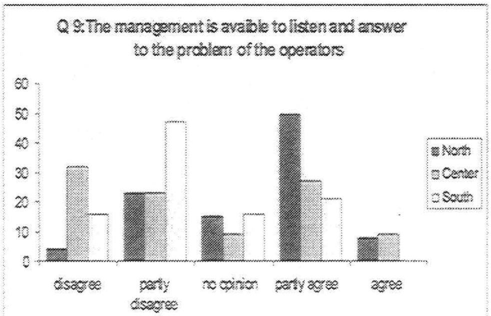


Chart 1: Evaluation of item 9 according to the different areas

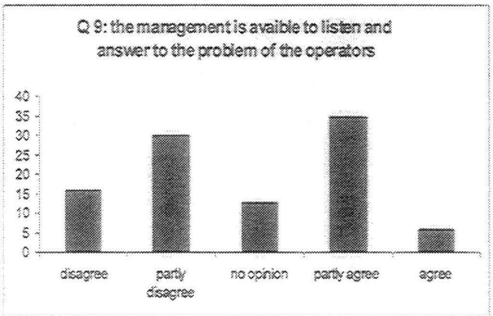


Chart 2: Evaluation of item 9 according to the entire sample

CONCLUSION

The study on the railway organisation has led to the definition of practical recommendations for reducing the risk of unsafe acts performed by operators, especially human error in the train driving task. The aim of the approach was to carry out a human factor review of drivers'safety performance in order to identify factors, which could influence a decrease performance and cause possible accidents and critical situations.

The analysis carried out by means of interviews and questionnaires, allowed to identify a number of important factors affecting the behaviour of operators and consequently to produce a series of recommendations for improving safety standards.

However, a specific outcome, though not initially planned, turned out to be particularly significant concerning the intra-cultural differences existing within the organisation. The results, while did not affect the overall (macro-level) set of recommendations on how to ameliorate the general safety standard of the company, may be very important at a second level (micro-level) of possible interventions that could aim at solving certain problems at regional level rather than at broader national level.

REFERENCES

Billings, C. E. (1997). *Aviation Automation: The Search for a Human-Centered Approach.* Lawrence Erlbaum Associates, Mahwah, New Jersey.

Degani, A., & Wiener, E. L. (1994). Philosophy, policies, procedures and practice: The four "P"s of flight deck operations. In N. Johnston, N. McDonald, and R. Fuller (Eds.). *Aviation Psychology in Practice,* Avebury Technical, Aldershot, UK. pp. 68-87.

Edwards, E. (1972). Man and machine: Systems for safety. In Proceedings of British Airline Pilots Association Technical Symposium. British Airline Pilots Association, London. pp. 21-36

Kirwan, B., & Ainsworth, L. K. (1992). *Guide to task analysis. Taylor and Francis,* London.

Redmil, F., & Rajan, J. (Eds.) (1997). *Human Factors in Safety-Critical Systems.* Butterworth, London, UK

Hawkins, F.H., (1975), *Human Factors in Flight,* Ashgate.

Sheridan, T. B. (1992). *Telerobotics, Automation and Human Supervisory Control.* The MIT Press, Cambridge, MA.

Wiener, E. L., & Nagel, D. C. (Eds.) (1988). *Human Factors in Aviation.* Academic Press, San Diego, CA.

RÉSUMÉ

LES DIFFÉRENCES CULTURELLES INTRA-NATIONALES RÉVÉLÉES PAR LES ANALYSES DES FACTEURS HUMAINS DANS LES SYSTÈMES SOCIO-TECHNIQUES COMPLEXES

Cette communication présente une étude, à la fois théorique et de terrain, dans une grande entreprise, visant à évaluer des facteurs importants qui affectent le comportement d'opérateurs et à établir des recommandations pour améliorer les normes de sécurité. De plus, dans la mesure où l'étude a couvert un large spectre de facteurs et de questions, un résultat inattendu a été obtenu : les différences en matière de culture de sécurité entre différentes régions d'un même pays. Ce résultat s'est avéré particulièrement significatif et pourrait conduire à repenser les politiques de sécurité dans l'organisation.

MOTS CLÉS : Données facteurs humains, Analyse des accidents, Études de terrain, Systèmes socio-techniques, Études organisationnelles.

CSAPC’99
AUTHORS /AUTEURS

RÉGION
NORD
PAS DE CALAIS

Conseil Général
Département du Nord

Villeneuve d'Ascq

TRANSTECH

Editors / Coordinateurs

Jean-Michel HOC

Patrick MILLOT

Erik HOLLNAGEL

Pietro Carlo CACCIABUE

ISSN 0760 5641

200 F TTC

Achevé d'imprimer par nos soins
Ⓒ Presses Universitaires de Valenciennes - 3ème trimestre 1999