REX: A HUMAN FACTOR FLIGHT SAFETY RESEARCH PROGRAM

N. Maille, L. Chaudron, P. Le Blaye
ONERA

J.-Y. Grau
IMASSA

Keywords : *human activity, aviation safety, flight analysis, cognitive modeling*

Abstract

ONERA and IMASSA are currently involved in a research program which aims at defining a methodology for the analysis of military aircrews' activity during everyday operations.

The proposed methodology is different from the flight analysis processes currently used by airlines, for instance, as it is clearly focused on the human activity : it is based on a model of generic activity, built from aircrews' interviews and flight simulations.

This model is used as a reference to identify possible discrepancies between the generic activity and the actions actually performed during the flight. Those discrepancies are then interpreted using a typology of typical safety relevant events, which was established through an extensive review of incidents cases.

1. Introduction

ONERA (*Office National d'Etudes et de Recherches Aérospatiales*) and IMASSA (*Institut de Médecine Aéronautique du Service de Santé des Armées*) are currently involved in a research program called REX-FH, which aims at defining a methodology for the safety analysis of military aircrews' activity during everyday operations, using the data recorded on board the most recent aircraft.

This research program is supported by the Human Sciences division of the French MOD (DGA/DSP/STTC/SH), for the need of the French Air Force.

A review of existing experience feedback systems and available methodologies for on-line activity analysis has been conducted during the first phase of the research program [1].

The second phase of the program aims at developing a methodology for the analysis of the actual activity of an aircrew through the mission data recording. The generic activity during this mission was described and modeled, together with safety relevant events to be identified [2].

The funding principles and the methodology developed under this program are described in this article. Some perspectives are also opened for the continuation of this research program.

2. Background

Whatever the domain (aviation, nuclear power plants, other transportation means or industries), the study of human operators activity at work inevitably reveals the extreme variability of human behavior.

The behavior of the operators is indeed the visible expression of the variability. The intra and inter individual variability primarily results from the various contextual adaptations of the operator's knowledge and know-how.

In the actual mission situation, similar results may be obtained by different procedures. The choice of a particular procedure is obviously not driven by a will to violate the prescribed rules, but by the necessity to save cognitive resources, to manage risks and eventually to keep the situation under control.

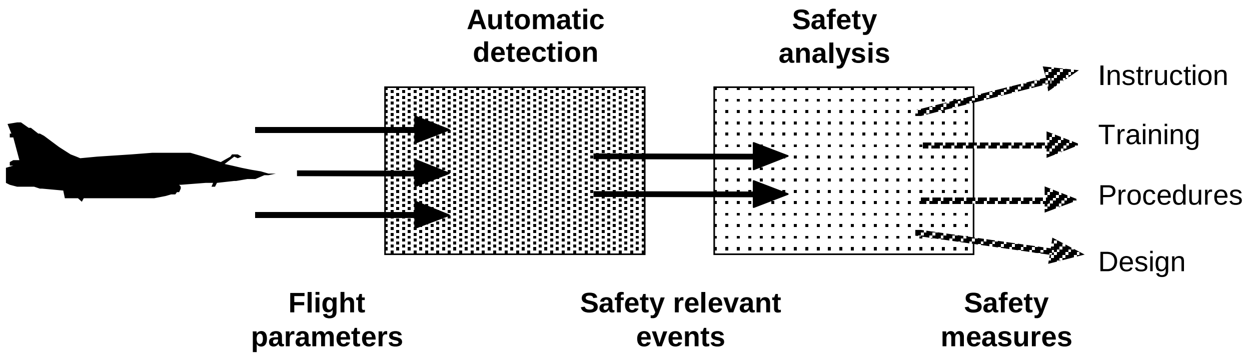

Figure 1 : the two steps process of flight analysis systems.

In order to improve the safety or performance level of socio-technological systems, the measures undertaken on instruction, on ergonomics or on professional rules and procedures should carefully address this variability.

In particular, appropriate methodologies are needed to get an objective feedback on the variability of aircrews' activity in everyday operations. It should be possible to collect and to analyze some relevant elements of knowledge and know-how in order to formulate the relevant recommendations for safety improvement.

The review of existing experience feedback systems carried out during the first phase of the program clearly showed that various systems already exist in most of the modern complex industries and transportation means, in order to better understand what are the difficulties actually faced by the operators in control of the processes.

Aviation has always given a particular attention to experience and "Human Factors" feedback. Mandatory or volunteer and confidential reporting systems have been used for a long time and will certainly develop as more accurate data mining techniques made them worth.

Those reporting systems have proved to be really useful, although their efficiency for the improvement of safety widely depends on the level of confidence they're accorded by the operators. Nevertheless, their efficiency is limited because of the subjective nature of the information provided and because of extra biases due to the associated regulation.

In order to get an accurate picture of the actual practices in flight, various tools are now emerging that tend to address not only the events revealed by a consecutive incident, but also less visible facts which could help understand what are the possible safety hazards of the socio-technological systems.

One of the most systematic and objective tools is the day-to-day analysis of the flight recorded parameters which is now set up by most airlines with the corresponding hierarchical structure. These Flight Operational Quality Assurance (FOQA) programs offer a substantial amount of objective information and their safety benefit is high, especially if the aircrew involved in an event is allowed and agrees to participate to the analysis process.

These flight analysis systems use a two steps process in order to detect and analyze well-known safety critical events (fig. 1). They are primarily based on the detection of deviations from flight profiles and exceedances of parameters limitations. The analysis of the detected events is then left to a skilled specialist.

The current tools look for deviations from a normative perspective of the crew activity : the reference consist of a prescribed task ; the search events are pre defined and their detection process doesn't take into account the possible different ways of achieving the mission goal.

Of course, those flight analysis systems provide useful data for safety improvement, but they're limited to deviations from prescribed rules and focused on “negative” events, which may have an immediate impact on flight safety.

New ways are now open to better know the actual practices of the crew : based on a representation of activity which integrates elements of intra- and extra-individual variability, an extended definition of the deviations in the aircrew behavior may be proposed and addressed in order to establish more accurate safety recommendations.

Our approach refers specifically to theoretical and experimental studies carried out by the department for Cognitive Sciences of IMASSA, concerning the cognitive aspects of aircrews' activity, error mechanisms and flight safety [3].

3. Objective

The objective of our research program is to develop a mock-up of a possible systematic flight analysis software for military aircraft that would be able to identify safety relevant events from a "human factors" perspective, using the available data recording of each mission. Those events will then be studied by a safety analyst who will formulate the appropriate safety recommendations.

The general purpose of the research program is to establish the necessary methodology to achieve this objective. The key of this methodology is to model the aircrews' activity and also the events which are thought to be safety relevant and that we intend to address.

Those events will be identified as gaps between the actual activity and a reference norm. The criteria for human variability and the interpretation of the gaps in terms of safety depend directly on the definition of this reference norm [4]. In the most classical approach, this norm may be defined as the prescribed task, as written in the documentation (legal requirements and recommended procedures). The norm may also consist of the usual practice, that may differ from one operator to another and also for one operator during his professional career. Last, the reference norm may be considered as the primary intent of the operator and a gap from this norm will be defined as a human error [5].

The choice of the reference norm has various implications from a safety point of view. Given the reference norm, it becomes possible to address not only the relatively well-known "negative" deviations and mistakes, but also the difficulties that the operators have to face in normal operations and the associated solving mechanisms that finally benefits to safety.

This approach is consistent with the ecological perspective of safety promoted by IMASSA. The error is understood as a natural element of human activity and it participates directly to the learning process, to the saving of cognitive resources, and so, to the regulation of the activity [6].

The aim of a safety approach is not to suppress the human errors but to avoid that they could lead to an accident. Rather more, it is believed that much safety benefits could be learned from a better understanding of the necessary adaptations of aircrews' activity faced to their actual mission constraints. That's the motivation to build an objective methodology to better know and understand what is actually done in the cockpit.

More precisely, within this approach of safety, the final objective of an automated flight analysis system is to be able to identify the following events :

- a procedure differing from the prescribed procedure ;
- a procedure differing from actual known procedures used by operators ;
- an adapted known procedure ;
- an exceedance of a limitation due to an incorrect application of a procedure ;
- an unknown procedure.

4. Methodology

4.1. Activity models

In our methodology, the process for the identification of safety relevant events is primarily based on the description and modeling of aircrews' activity (Fig. 2).

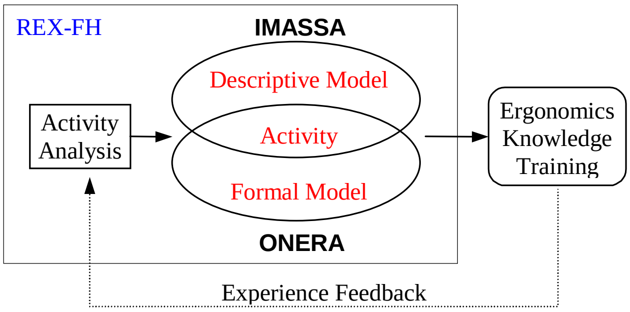

Figure 2 : Activity Models developed in REX.

This description was established on the basis of information directly collected from aircrews by IMASSA. It is called the Activity Descriptive Model (ADM).

It integrates a typology of Safety Relevant Events (SRE), which is used to guide the detection process.

These expertise elements are implemented using an appropriate formalism and an advanced programming language, resulting in an Activity Formal Model (AFM).

The methodology used to build those activity models is described in further details in the following of this paper.

4.2. Available objective mission data

The implemented model of activity is used as a reference frame to rebuild the activity actually performed, through the flight parameters recorded on board the aircraft during the mission.

Modern digital flight data recorders allow the collection of several hundreds of parameters at a time rate which is generally high enough compared to the dynamics of human decision making and actions.

Those parameters provide objective data on the aircraft state and parameters, but also on the systems currently in use, on the crew actions (position of most controls and switches) and on the warnings delivered to the crew.

Given the almost unlimited capacity of these new recorders, the assumption can be made that both the exact state of the aircraft systems and all the physical actions of the aircrew can be observed after the mission.

If they are considered alone, with no supplementary information about the mission, numerical data are of limited value when trying to analyze human intents and performance. This fact reinforced the requirement for detailed and faithful activity models to support the analysis.

In the context of post mission analysis, the problem of intent recognition is simplified because all data are available during the whole mission time. However, the assumption still has to be made that an observed achieved goal actually corresponds to the initial intent of the aircrew. In other words, human error defined by Reason as deviation from intention may be hardly observed through numerical data, unless all possible intents are captured in the activity model ; this objective is probably achievable when considering routine highly procedural flight phases such as take off or landing.

Other types of mission data are available on modern combat aircraft. The rapid advances in information processing will probably facilitate their interpretation by automated systems in the near future.

Audio records of voice communications are available and speech recognition system are already on board recent aircraft for the control of basics aircraft functions or information requests. The analysis of voice recorders could be very helpful, especially in the context of two seats aircraft –as seen during the experiment conducted within REX- where critical information is exchanged between the pilot and the weapon system officer, including check list and vital actions. These voice exchanges are usually formalized under recommended operating procedures, which could made their automated recognition quite easier as natural language.

Video records are also already available for the need of mission debriefs and weapon imagery. Their use could be helpful for a better understanding of decisions related to meteorological conditions or tactics (formation flight for instance).

Video records of the pilot’s external view including the head-up display were used to support the debriefings of the REX simulation trial, but the possible automated use of this longer term capability has not been addressed yet.

4.3. Identification of Safety Relevant Events (SRE).

This first step of the analysis of aircrews' activity consisted in a study of incidents reports from the French Air Force.

Those reports only concern incidents which are not submitted to a detailed investigation. Given the current experience feedback practices in the Air Force, they represent the closest available information to what could be collected with a systematic analysis system.

The study of 615 reports over the period 1991-1998 showed that the most safety critical phases of the mission is landing -average 22 % of the total reports- and take-off, in a much smaller extent.

With the agreement of the Air Force, it was decided to first focus the study on the landing phase, as the example application phase.

Among the reports concerning landing incidents, about half were identified as primarily involving human factors.

A more detailed analysis of the reports showed that the incidental scenarios could be categorized with a typology of the incident deviations, derived from the phenotypic model of error production [7]. The erroneous action is described here as a deviation from what should have been done, which is formally described outside of the actual situation, and generally based on the published procedures. The proposed typology is derived from an existing error categorization [8] with the adjunction of a temporal reference to better address the possible failures in time management [9].

The typology consists of the four following types of deviations, which have been indeed identified among the reports :

- intrusion (*e.g. during approach, fuel tank dropping instead of landing gear down*)
- inversion (*e.g. explosion of a tyre at landing after a failure of the anti-slide system, because the pilot breaks before using the parachute, although the procedure is to use the parachute first*).
- omission (*e.g. the crew forget to report his position to the control before starting a descent*)
- erroneous time management, which can be further divided under four types :
  - ∗ too early (*e.g. landing gear down while speed still too high*)
  - ∗ too late (*e.g. thrust reduction too late, which results in a precipitated approach)*
  - ∗ too long (*e.g. thrust kept idle for a too long time, inducing an engine extinction*)
  - ∗ too short (*e.g. landing flare too short, causing a hard touch down*)

Those deviations will be applied to the components of the descriptive model which is presented below.

4.4. The Activity Descriptive Model (ADM).

The second step of the analysis of the activity consisted of detailed interviews of 5 qualified military crew members, on the basis of paper scenarios.

The aim of these interviews was to capture the knowledge and the know-how of the crews for the landing phase. They were conducted by IMASSA on the basis of paper scenarios designed by an instructor pilot and similar to air task orders given to crews when they are assigned a mission.

A complete transcription of the interviews was systematically achieved and used to build the descriptive model of the activity.

This ADM consists of a generic structure, based on the schema approach [10]. This approach is now well accepted in cognitive psychology and also used for knowledge representation in computer science.

The ADM allows to represent the generic activity as it is described by the operators. It is as close as possible to the actual known practices and to the mental representation that the operators usually have after they have prepared their mission.

It integrates not only the most usual way of conducting the flight phase but also the alternative procedures envisaged by the crew, in order to face possible in-flight mission replanning, failures or context changes.

The mission phase is decomposed under elements corresponding to the various level of representation of the mission task, with the following typology, from the higher level to the most detailed :

Phase > Schema > Module > Cue

The Phase is defined as the highest level of description of the chronological follow on of the mission (e.g. "Landing"). A phase consists in a sequence of schemas ; for instance, the usual “landing” phase is composed of 8 schemas.

The Schema is a description of a particular goal-oriented activity (e.g. "Prepare landing configuration”, Fig. 3).

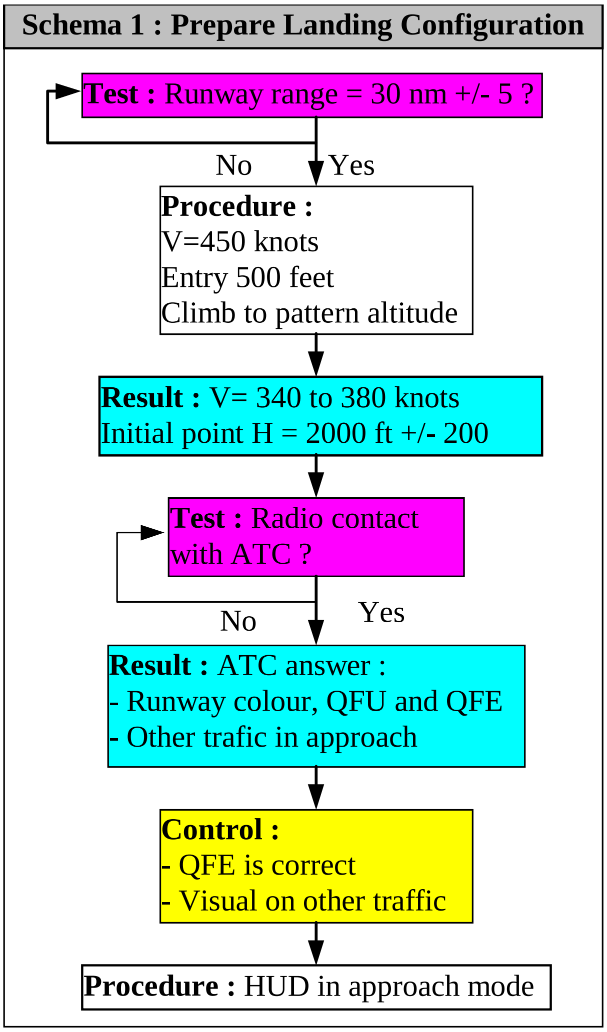

Figure 3 : an example schema script from the ADM approach mode

A schema has the following components :

- some entry conditions ;
- an initial state ;
- a script which is an oriented graph of modules, as shown in the example below ;
- a final state ;
- some exit conditions.

The Module is a typed consistent combination of cues. Each module relates to a particular cognitive mechanism : its type can be either a Test, a Procedure, a Result or a Control (e.g. Procedure : {Speed 450 knots, Intercept 500 feet, Climb to pattern altitude}).

A Cue is a piece of information, a property or an elementary physical action (e.g. "Speed 450 knots").

The MDA also integrates a description of the margins allowed on the value of key parameters of the current activity.

These margins are described as values of the expected precision of some Cues (e.g. “V = 340 to 380 knots”). The margins usually refers to the precision with which the corresponding flight parameters have to be controlled so as to guarantee a sufficient level of flight safety.

The value of the margins can sometimes be given by the aircrew with a certain indetermination, but most of the time they’re better defined by operational experts such as instructors.

Other margins also have to be integrated when the ADM is confronted to actual flight data profiles. These margins introduced by the analysis system user on the basis of his own knowledge of the activity should be clearly identified as such.

Although this last artefact is necessary to cope with the rigidity of numerical exceedances detection algorithms (rather than with the normal relatively imprecise control of human controlled processes), we think that the representation of these margins and of their possible dynamic regulation according to the context are key features to better address the variability of human activity.

4.5. Experimental data collection : the REX simulation trial

In a last phase of the analysis of crew activity, some man-in-the-loop mission simulations were conducted at the Istres Flight Test Center, using a realistic full-size simulator of a modern two-seats fighter aircraft (Dassault Mirage 2000D).

This simulation trial involved an operational military aircrew, actually flying on this aircraft type. An experienced instructor pilot controlled the events introduced in the scenarios ; he also evaluated the performance of the aircrew from an “academic” perspective.

Four mission scenarios of one hour each were flown, including specific events (various technical failures, meteorological changes and in-flight rerouting) in order to collect actual examples of deviations.

Detailed debriefings were conducted with the crew members, using the video recording of the pilot’s external view including the head-up display, in order to support the activity analysis.

Numerous flight parameters were of course recorded during the simulations, to be used for the development of the software mock-up.

The data collected during this simulation trial were exploited to validate the MDA and the typology of SRE.

Some example cases of deviations appeared during the simulated missions, such as hurried approaches or an exceedance of a speed limitation imposed by the engine calculator failure.

The scenarios also help to identify some particular unknown practices that had not been explicitly described by the crews during the previous interviews.

For instance, different final approach procedures (including different values of the target angle of attack and different choices of control parameters) were described by the aircrew and by the instructor ; this example case of procedure adaptation confirmed, if necessary, the existing variability of practices and know-how, even between aircrews from the same squadron.

4.6. Creating the Activity Formal Model (AFM).

The AFM is devoted to mathematically capture the ADM under a rigorous and generic formalism.

It is designed to be easily implemented, using an appropriate programming language and including the necessary software functionalities associated with flight data processing.

The requirements of the AFM are:

1) to provide a formal and computable representation of the activity characteristics as revealed and validated by the ADM,

2) to ensure the data connections between the elements of the activity and their corresponding effects on the flight parameters recorded by the on-board devices.

Indeed, a discrepancy feature which was assessed in the ADM (as it appeared relevant from a human factors analysis) must be consistently represented as a graph operator in the AFM. For instance, an ADM inversion of two Cues observed in a given activity should be captured by a specific permutation in the AFM, and the application of this operator should be confirmed by the correlation with the flight parameters.

From a microscopic point of view, the informal notion of deviation from the activity described in the ADM has got a corresponding definition as a structured difference between two pieces of knowledge.

Thus, the AFM aims at providing formal definitions of naives notions, such as the expression "addition of errors" which is frequently described by accidents investigators. Is the naive denotation "addition" formally relevant? Should it be considered that this "addition" is commutative, associative or distributive with other operators? The purpose of the AFM is to filter, to structure and to assess the amorphous knowledge in a consistent and relevant way.

From the theoretical point of view, the Cues are coded as constrained litterals of the first-order logic (e.g. "speed(450)"), they represent the smallest piece of knowledge on which validity questions can be performed.

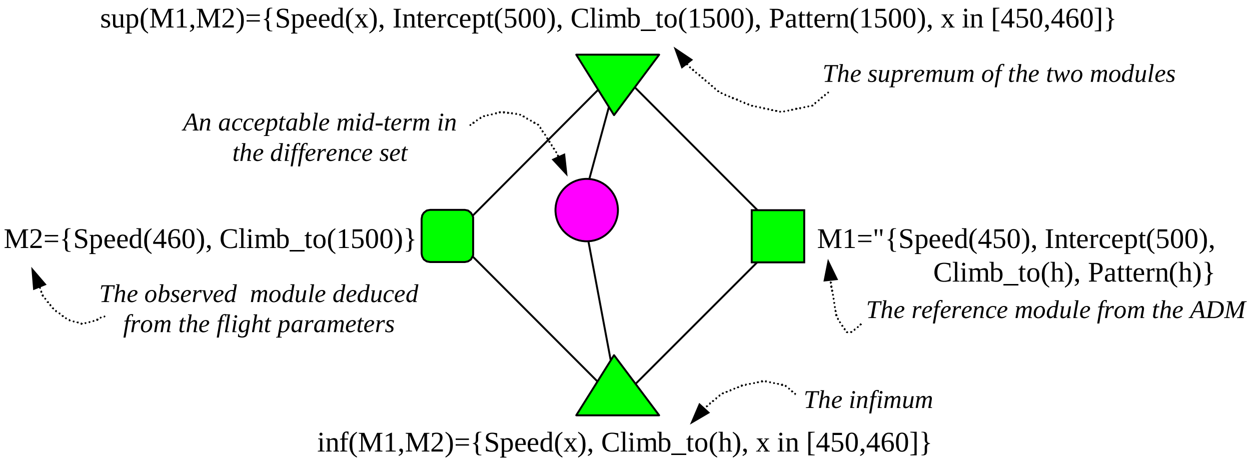

Figure 4 : An example application of modules matching using the Cube

Hence, the Modules are constituted of conjunctions of litterals (e.g. M1={Speed(450), Intercept(500), Climb_to(h), Pattern(h)}).

An algebraic framework, called the Cubes model, is used to completely describe and capture the differences between pieces of knowledge such as different Modules.

We think that such a skill is the key of activity deviation modeling. We previously established that the Cube model is a complete non-modular lattice [11] ; all the symbolic fusion operators (supremum, infimum, etc.) are defined in the model and implemented in Constrained Logic Programming (CLP), more precisely in PROLOG.

The Cube model provide the necessary functionalities to the general problem of partial matching of symbolic information items.

For instance, considering a second Module M2={Speed(460), Climb_to(1500)} deduced from the flight parameters of an actual mission, the Cubes model makes it possible to compare M2 to M1, and to determine the symbolic difference set of these two pieces of knowledge (Fig. 4).

The landing phase has been completely formalized and implemented. The main functionalities have been demonstrated on some example cases using actual data from the simulation trial.

The demonstrated functionalities include the detection of inversion, omission and insertion of cues or modules, consistently with the proposed taxonomy of safety relevant events.

The current research on AFM focuses on the integration of constraints into the Cubes model so as to fit closely to the actual expressions of both flight parameters and the refined modules of the ADM.

5. Possible extensions and conclusion

The first phases of this research program demonstrates the feasibility of a systematic flight analysis taking aspects of human variability into account.

Several research work are still required to reinforce the accuracy and the robustness of the proposed generic model of activity. Experiments using actual mission data should be conducted in order to address real world practices and to validate the detection margins.

Theoretical development and knowledge acquisition are expected to extend the functionalities of the proposed system related to the interpretation of the differences identified between the generic model and the actual activity. For instance, the typology of safety relevant events should integrate elements of expertise collected from flight safety officers.

Research is also needed in order to better understand and integrate the dynamic adaptation of the safety margins depending of various factors such as, for instance, the mission context or the pilot’s level of expertise [12].

Based on a careful analysis and generic modeling of the crew activity, our approach should contribute to get a better knowledge of the variability of everyday practices, to better understand how the crews actually adapt their procedures to face the operational constraints and, eventually, how to avoid incidents becoming accidents.

6. References :

[1] Le Blaye P., Doniat C (1996). Recensement des besoins et des méthodes de retour d'expérience sur les facteurs humains pour la sécurité de vol des aéronefs militaires. *ONERA Technical Report 3/5276 SY*, 1996.

[2] Chaudron L., Grau J.Y., Le Blaye P., Maille N. Retour d’expérience facteurs humains pour la sécurité des vols. REX_FH phase 2. *ONERA Final Report. AER.00710.01F*, May 2000.

[3] Amalberti R. La conduite des systèmes complexes. *Collection Le Travail Humain*. Editions puf. 1996.

[4] Grau M.P., Bellet S.C. Bases théoriques de l'erreur humaine en situation de pilotage. *IMASSA Report n°93-27*, Dec. 1992.

[5] Reason J. *Human error*. Cambridge University Press, 1990.

[6] Amalberti R., Wioland L. Human error in aviation. *Aviation safety conference*. Rotterdam, Aug. 1997.

[7] Hollnagel E. The phenotype of erroneous actions. *Int. J. of Man-Machine Studies*, 39, 1-32, 1993.

[8] Swain, D. and Guttmann, H.E.. *Handbook of Reliability Analysis with emphasis on Nuclear Plants Applications*. Nuclear REGulatory Commission, NUREG/CR-1278, Washington D.C., 1983

[9] Decortis, F and Cacciabue, P. Temporal dimensions in cognitive models. *4th IEEE conference on Human Factors and Power Plants*, June 5-9, Monterey, CA, 1988.

[10] Schank R.C., Abelson R.. *Scripts, plans, goals and understanding*. Hillsdale, N.J. Lawrence Erlbaum Associates Inc., 1977.

[11] Chaudron L., Cossart, C., Maille N. and Tessier C. A purely symbolic model for dynamic scene interpretation. *International Journal on Artificial Intelligence Tools*, 6(4):635-664, 1997.

[12] Wioland L. and Amalberti R. Human error management : towards an ecological safety model. *9th European Conference on Cognitive Ergonomics*. Limerick University, Ireland., 1998