Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: http://oatao.univ-toulouse.fr/
Eprints ID: 11686

To cite this document: Lacko, Ivan and Moravek, Zdenek and Osterloh, Jan Patrick and Rister, Frank and Dehais, Frédéric and Scannella, Sébastien *Modeling approach to multi-agent system of human and machine agents: Application in design of early experiments for novel aeronautics systems.* (2013) In: 11th IEEE International Conference on Industrial Informatics (INDIN) 2013, 29 July 2013 - 31 July 2013 (Bochum, Germany).

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@inp-toulouse.fr

Modeling approach to multi-agent system of human and machine agents: Application in design of early experiments for novel aeronautics systems

Ivan Lacko, Zdenek Moravek
Advanced Technology Europe
Honeywell International s.r.o.
Turanka 100, Brno 62700,Czech Republic
{Ivan.Lacko, Zdenek.Moravek}@honeywell.com

Jan-Patrick Osterloh
OFFIS Institute for Information Technology
Escherweg 2, 26121 Oldenburg, Germany
osterloh@offis.de

Frank Rister
truestream aerospace human factors
Alphonsstrasse 15
22043 Hamburg, Germany
frank.rister@hamburg.de

Frederic Dehais, Sebastien Scannella
Aeronautic and Spatial Center (CAS – ISAE)
10 av. E. Belin, 31055 Toulouse Cedex 4, France
{Frederic.Dehais, s.scannella}@isae.fr

Abstract—Design of future systems for flight-deck automation will reflect a trend of changing the paradigm of human-computer interaction from the master (human)-slave (machine) mode to more equilibrated cooperation. In many cases such cooperation considers several humans and computer systems, for which multi-agent dynamic cooperative systems are appropriate models. Development of such systems requires very profound analysis of mutual interactions and conflicts that may arise in such systems. Additional testing is exhaustive and expensive for such systems. In the scope of the D3CoS project these problems are addressed from the modelling point of view with ambition to create tools that will simplify the development phase and replace parts of the testing phase. In this paper we investigate common flight procedures, for which computer assistance could be developed. We show how formal modelling of procedures allows us to inspect procedural inconsistencies and workload peaks before the development starts. We show how a computer cognitive architecture (a virtual pilot) can simulate human pilot behaviour in the cockpit to address questions typical for the early phase of the development. Analysis of these questions allows us to reduce the number of candidates for the final implementation without the need of expensive experiments with human pilots. This modelling approach is demonstrated on experiments undertaken both with human pilots and a virtual pilot. The quality of the outcome from both experimental settings remains conserved as shown by physiological assessment of pilot workload, which in turn justifies the use of the modelling approach for this type of problems.

Keywords—system design; flight-deck automation; model-based approach, cognitive architecture

I. INTRODUCTION

Standard operating procedures (SOP) describe pilots' activities in the cockpit during flight. They are defined to optimize workload of a pilot when applied appropriately to a flight situation. It happens often that the current situation changes because of flow optimization in air traffic management (ATM) and subsequently implies a pilot to change the strategy he has planned. A correct reaction of the pilot is to estimate the current state and to evaluate available procedures to select the most relevant one. The situation is therefore accompanied with an increase of workload and becomes prone to errors.

The development in aviation offers tools to support pilots in such situations with an ambition to supply relevant information in a form that reduces the overall workload and facilitates the decisions a pilot makes. The amount and form of displayed information is crucial in stress situations and a common way of handling the problem is to create several alternatives that are later probed in experiments with pilots. The experiments are, however, expensive and time consuming. We therefore discuss in this paper another approach – the model-driven approach:

1. A symbolic language is used to describe the SOPs. Using a processor of this language, each procedure can be parameterized for cockpit geometry so that reaction/movement times and workload can be derived from psycho-physiology models.
2. Cognitive models (a virtual pilot) are used to simulate pilot's behaviour in computer experiments to complete the set of parameters needed to evaluate alternative designs.

Such an approach is appealing for time/cost reduction and a potential for fast adjustment of experiments to new or updated requirements. Another benefit is its applicability to any phase

of the development, including the early phases of concept definition.

On the other hand, the equivalence of information obtained from human and computer experiments needs to be justified. Therefore we design an experiment to investigate the ability of a virtual pilot to replace human pilots in evaluation of cockpit tools. In this paper, we will describe therefore both aspects of this approach, first the model-driven approach, and second a first evaluation of the model-driven approach.

II. METHODS

A. Demonstration use-case

A simple use-case of position selection will be used in this paper. The cockpit geometry allows a tool to be placed at several alternative positions. However, each position may require a different effort for certification and also may offer a different support to a pilot.

For example, a new tool for a primary flight display (PFD), which is in the straight view of a pilot, will be most supporting, but it will require a bigger effort to certificate such a tool. An electronic flight bag (EFB) application on the other hand may be easy to certificate (as it has less access to primary flight deck systems), but pilots will have to divert their focus to read its information, which may not be possible in high workload situation.

Therefore, a compromise between development/certification effort and availability must be analysed before going for a concept. This use-case can be easily assessed in a computer experiment to demonstrate the potential of the model-driven approach.

B. SOP modelling

Standard Operation Procedures (SOPs) are modelled as a set of rules, which define a task to be performed. The rule language is based on so called Goal-State-Mean (GSM) rules [2]. Each GSM rule contains a left and a right side; see also Fig 1 for an example rule.

The left-hand side defines the conditional element which contains the Goal- and the State-Part. Each rule is responsible for exactly one goal, and can be chosen if the Boolean expression of a condition is true. The right-hand side specifies a method (an action to be started) expressed in the Means-Part of a rule. The actions are defined to account for memory, perception and motor operations. Additionally, the language is able to specify different temporal relations giving the order and priorities (likelihoods) of the sub-goals [1].

For example, the rule in Fig. 1 can be informally be read as "IF the actual goal is to retract the gear and the aircraft has lifted off, THEN pull the gear lever, shift attention to gear annunciation, pursue the goal to check if the gear actually retracts and afterwards pursue the goal to call out the gear state. This rule defines a goal-subgoal relation between GEAR_UP and subgoals CHECK_GEAR_UP,

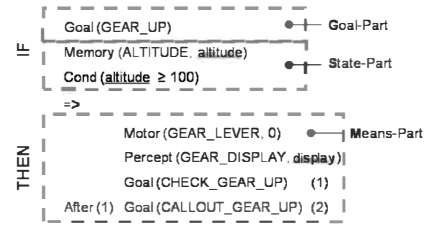CALLOUT_GEAR_UP. Between the subgoals a temporal order is imposed (by "After").

Fig. 1: Structure of a rule for SOP modelling

For modelling of the SOPs, a tool that allows graphical modelling of such rules has been developed, called PED (Procedure EDitor).

C. Cognitive architecture

The cognitive architecture CASCaS, used in our model-based approach has been developed by OFFIS since 2004, see for example [1], [2], [3], [4]. The cognitive architecture is based on the theory of Anderson's three behaviour levels [5]: autonomous layer (tasks processed without thinking in daily operations), associative layer (tasks processed by formal procedures) and cognitive layer (tasks requiring new plans in unfamiliar situations), see Fig 2. An additional layer stands for perception and motor behaviour as an interface to a simulated environment. The procedures on the associative layer are stored in the memory component in form of the previously described GSM rules. At each moment, system state and processing of a procedure create the mental model, and are expressed as an ordered set of goals and subgoals that have to be accomplished – the so called goal agenda.
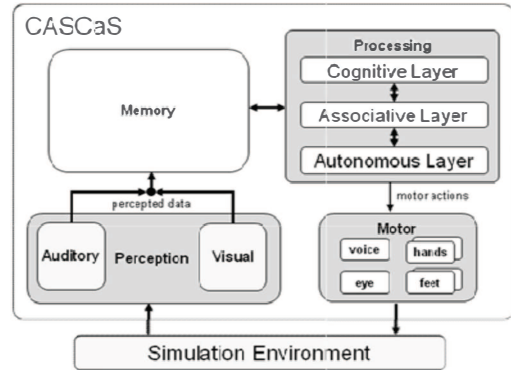

Fig. 2: Layered Cognitive Architecture

Processing of the goal agenda follows these steps:

1. A goal is randomly selected from the goal agenda, which contains all currently selectable goals.
2. All rules containing the goal in their Goal-Part are collected, evaluated and hierarchically organized into a rule set.
3. One rule is randomly selected from the rule set and fired, which means that the motor and percept actions are sent to the motor and percept component

respectively and the subgoals are added to the goal agenda.

This process is iterated until no more rules are applicable. For the prediction of task execution times, each component has a psychological and physiologically sound model of timing and behaviour, i.e. in the memory component the remembering and forgetting of information is modelled, or in the eye-motor component sophisticated models of eye-movements are integrated. In order to take the positions of the instruments for the calculation of the eye-movements into account, the locations of all instruments are modelled based on the original cockpit layout, and loaded into CASCaS, see also Fig. 3 for a visualisation of that input. For the computer experiment, the advisory system can then be re-placed by simply changing the model of the cockpit (XML-Format).

Fig. 4: B737 Layout within CASCaS; Yellow dot marks the current eye-position of the model during simulation (on the leftmost instrument, the EFB)

Previous work has shown, that CASCaS can be used to predict gaze and eye behaviour [3].

D. Experiment design

There are two experiments presented to demonstrate the benefits of the model-driven approach. First, experiment with human pilots assesses the benefits of a supporting tool itself. Second, a computer experiment evaluates various placement alternatives of the tool in the cockpit. Both experiments use defined SOPs for a regular approach.

There are two types of outcome from the experiment – those related to the supporting tool and those related to applicability of the computer experiment.

In the experiment with human pilots, stress situations are induced during the approach by unexpected situations such as device failure or a new clearance (shortening of available distance to the runway). A supporting tool is placed at the PFD and the following aspects are monitored in pilot behaviour:

1. Time series of relative workload
2. Pilot's gaze ratio for looking at the supporting tool
3. Pilot's consideration of information provided by the supporting tool

4. Relation among the relative workload, pilot's action, quality of information from the supporting tool and usage of that information.

The experiment was undertaken in a B 737 NG full-flight simulator at Lufthansa Training Centre, Berlin, with 26 airline pilots. Each pilot flew four different scenarios the simulator. Due to legal restrictions, we could place the advisory system only into one position (lower engine display, between both pilots). The subjects have been equipped with an eye-tracker as well as ECG measurement. The relative workload is derived from ECG measurements (heart rate variability). Gaze ratio is derived from eye-tracker measurements (head-mounted eye-tracker) and pilot's consideration of the information is evaluated from time correlation of pilot's reactions and displayed information.

Subjective measures of the situations were collected through NASA-TLX (Task Load indeX) questionnaires ([6], [1]).

The computer experiment evaluates accessability of information from the supporting tool. The tool is placed subsequently at four different locations. As benchmark, PDF is used, then Enhanced Ground Primary Warning System display – EGPWS, EFB on right of the pilot (in place of back-up display) and on left of the pilot (usual EFB position) are used. At each location, CASCaS follows the SOPs and estimates the accessibility with respect of time the pilot needs to divert his attention from the regular procedure to read the information from the supporting tool.

III. RESULTS

The analysis of the TLX questionnaires confirmed that pilots considered the events selected to induce the re-evaluation of plans as stressful, see Fig. 3. Relative workload analysis showed an increase of the workload value shortly after the event happened, see Fig. 4.
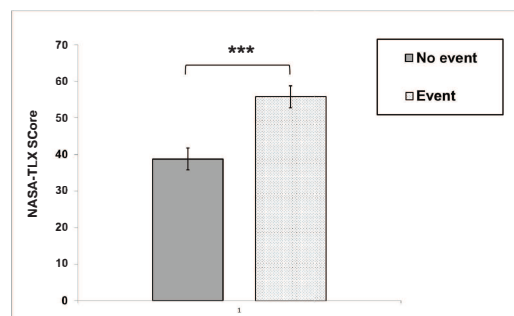

Fig. 3: NASA-TLX score for scenarios without induced replanning and with replanning; *Error* bars represent the standard error of the mean, *** $p<.001$

There is an obvious correlation for ECG data, eye-tracker data and event positions. The time series of relative workload derived from physiology measurements and from the computer analysis of SOPs correspond one to each other.

IBI
= 60000/
Heart_rate

LowFreq/HighFreq
*increase with mental
work load*

Eye tracking ratio
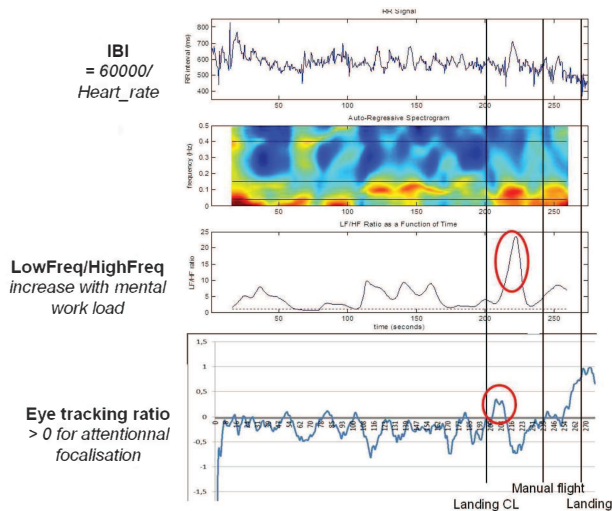*> 0 for attentionnal
focalisation*

Fig. 4: ECG, eye-tracker and event correlations

Additionally, the analysis of the ECG data shows the supporting tool reduces the overall workload during the approach, see Fig 5.
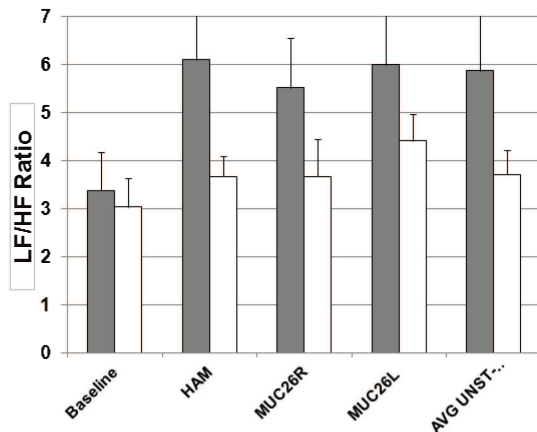

Fig. 5: ECG based workload estimates. White - with the supporting tool, gray without the tool.

The reactions of pilots with respect to the supporting information are evaluated on basis of NASA-TLX questionnaires. Preliminary results show generally good acceptance of information provided at the PFD. The investigation of causality between supporting information and pilot behaviour is now processing data from flight logs and eye-tracker data.

Having confirmed the benefits of the supporting tool in the experiment with human pilots, the next question is to evaluate whether the instructions provided by the tool can still be beneficial if placed at different locations than PFD. The answer has implications with respect to the restrictions that apply to the development of the tool.

The computer analysis of other positions for the supporting tool shows a decrease in efficiency of the displayed

information in order: EGPWS (12.33 s/cycle), EFB on right (12.82 s/cycle), EFB on left (13.47 s/cycle). At peak workload situations, time restrictions inhibit pilots to make use of EFB positions: the diversion of focus requires more time than available for a relevant procedure step. This conclusion was also confirmed by a simple test in the cockpit with several human pilots (out of scope of the experiment).

IV. Discussion

The preliminary data analysis suggests there are correlations between the workload predictions made by the computer version and workload data obtained from the human version of the experiment. Analysis of NASA-TLX questionnaires is in agreement with the data from the physiology analysis. All this justifies the use of a model-driven approach to evaluate display alternatives (or position alternatives in this study); especially in an early phase of development, when human pilot experiments are difficult to design. The on-going data analysis should allow deeper analysis with respect to improve the efficiency of the supporting tool.

The computer experiment proved to be useful in analysis of cockpit locations. It showed some of the locations should not be considered for future prototype development. The overall expanses and time requirements to set-up the computer experiment were significantly reduced (in about 20% lower). The future improvement of the modelling tools within D3CoS project will allow to deploy them in the more complex experiment (such as the human pilots experiment described here) to see further benefits from the model-driven approach.

V. References

[1] A. Lüdtke et al, "Modeling Pilot and Driver Behavior for Human Error Simulation", HCI 2009.
[2] A. Lüdtke, C. Möbus, „A Cognitive Pilot Model to Predict Learned Carelessness for System Design". In A. Pritchett and A. Jackson (eds). Proceedings of the International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero), 29.09.-01.10.2004, Toulouse, France: Cépaduès-Editions. CD-ROM, 2004
[3] A. Lüdtke, F. Frische, J.-P. Osterloh, Validation of a Digital Human Model for Predicting Flight Crew-Aircraft Cockpit Interaction. In Proceedings of 9. Berliner Werkstatt für Mensch-Maschine-Systeme 2011, VDI Verlag
[4] A. Lüdtke, J.-P. Osterloh, T. Mioch, F. Rister, R. Looije, "Cognitive Modelling of Pilot Errors and Error Recovery in Flight Management Tasks". In Proceedings of the 7th Int. Working Conference on Human Error, Safety, and System Development (HESSD) Brussels, Belgium, Sept. 23-25, 2009
[5] J. Anderson, Learning and Memory. John Wiley & Sons, Inc. (2000).
[6] M.M. Sepehr, "Assessment of Subjective Mental Workload Using NASA-Task Load Index". In Proceedings of the 7th European Annual Conference on Human Decision Making and Manual Control , 69-75. 1988
[7] S. G. Hart, L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research". In P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload. Amsterdam: North Holland Press. 1988