From qualitative assessment to qualitative accreditation with cognitive agents

Anton Minko, Eric Gouarderes

From qualitative assessment to qualitative accreditation with cognitive agents

Anton MINKO[1], Guy GOUARDERES[2]

[1]SARL « Interactive STAR », Centre Condorcet Développement 162, av. du Dr.A.Schweitzer 33600 PESSAC, France
[2]Equipe ISIHM - LIUPPA - IUT de Bayonne, 64100 Bayonne, France
e-mail : [1]anton.minko@star-ima.com, [2]Guy.Gouarderes@iutbay.univ-pau.fr

Résumé

Cet article présente les résultats d'application des modèles cognitifs à la formation aéronautique à travers l'utilisation d'un STI (Système Tuteur Intelligent) multi-agents. Plus particulièrement, l'article traite des modèles de l'erreur humaine et de l'application des technologies multi-agents afin de réaliser un diagnostic de ces erreurs et des écarts cognitifs associés. Le modèle de raisonnement basé sur la simulation qualitative fournit une large variété de paramètres constituant la base de l'évaluation pédagogique de l'apprenant. Un STI basé sur la simulation et utilisant l'approche "apprendre en faisant des erreurs" a servi de cadre pour les expérimentations. La démarche globale est supposée être utilisée en perspective d'e-accréditation de la formation, qui semble être inévitable dans le contexte actuel de globalisation et de développement du e-learning dans les compagnies aéronautiques.

Mots-clés: formation continue basée sur la simulation, agents et modèles cognitifs, simulation qualitative, systèmes multi-agents, e-learning

Abstract

This paper presents the results of application of cognitive models to aeronautic training through the usage of a multi-agent based ITS (Intelligent Tutoring Systems). More particularly, the paper deals with models of human error and application of multi-agent technologies to diagnose human errors and underlying cognitive gaps. The model of reasoning based on qualitative simulation supplies a wide variety of parameters as the base for pedagogical evaluation of the trainee. The experimental framework is simulation-based ITS, which uses a «learning by doing errors» approach. The overall process is intended to be used in the perspective of e-accreditation of training, which seems to become unavoidable in the context of globalisation and development of e-learning in aeronautic companies.

Keywords: simulation-based vocational training, cognitive models and agents, qualitative simulation, multi-agent systems, e-learning

Introduction

In the world of aeronautical training, many training tasks are more and more performed in simulators. Aeronautical simulators are very powerful training tools which allow to reach a very high degree of realism (perception of the simulator as a real aircraft by trainee). However, several problems may appear. One of the most critical problems is taking into account the behaviour of the trainee, which remains relatively limited because of the lack of the online feedback on the users' behaviour.

Our research is centred on the description and the qualification of various types of behaviours in critical situations (resolution of a problem under risk constraints) depending on committed errors. We articulated these issues using an adapted version of the ACT-R/PM model [1]. The first, fairly general source of errors in ACT-R models, is the failure of retrieval or mis-retrieval of various pieces of knowledge (in CBT, Computer-Based Training, systems – checked Flow-Charts, or in PFC – Procedures Follow-Up Component – in terms of ASIMIL, see end of paragraph). The second and more systematic error source is the time/accuracy trade-off in decision-making.

There are also other secondary sources of error, such as the trainee failing to see a necessary sign/indicator in the time provided in order to perform the needed operation. These sources of error are mainly due to ergonomics or psychology affects.

In this work we try to translate all of the above-mentioned sources/parameters of errors into triggered hallmarks that are in the learner profile [2]. We have considered a number and the possible extensions of the error types after Rasmussen's framework [3] and performed partial in-depth analyses about: level of reflexes (sensor-motor ability), level of rule-based errors (widely revised in aeronautic research [4]), level of trainee's cognitive abilities based on John Self's [2] theory about learner profile.

Our idea consisted in proposing a multi-agent system including revisable competencies of a human tutor in the framework of Actors-like agents architecture (autonomous and concurrent), where different agents are specialised in their respective types of errors.

This research was undertaken within the framework of project ASIMIL (Aero user friendly SIMulation based dIstance Learning) financed by the 5th Framework Program of of the European Community. The main objective of ASIMIL project consisted in exploring new approaches in the aeronautical training, including distance training, simulation, technologies of intelligent agents and virtual reality. The final prototype represents a real desktop simulator installed on a workstation over the network.

Cognitive modelling of training process

The general question raised in this paper: how to ensure a good quality of computer-assisted training equivalent or higher than that obtained in the classical training. We found the answer in the use of ITS and in the modification of the conventional training loop [5] by introducing cognitive models.

By definition, ITS is an adaptive system. Its adaptation is carried out via the modification of the internal representation of learning recorded and used by the ITS (learner profile). The system must build a personalised model of learner, allowing to adapt the course curriculum to the trainee, to help him/her browse the course and to carry out exercises, by providing personalised help.

Cognitive models provide the means of applying psychology to predict time, errors, and other measures of user interaction [6].

This leads us to restore the role of the instructor in the training, because the main task of computerisation in the training consists in returning to the instructor all the freedom of diagnosis and decision by decreasing his/her logistics' tasks for the profit of teaching. Moreover, the ITS is obliged to interact with all the components present in the conventional loop of training (trainee, trainer, simulator, training procedures) (see Figure 1).

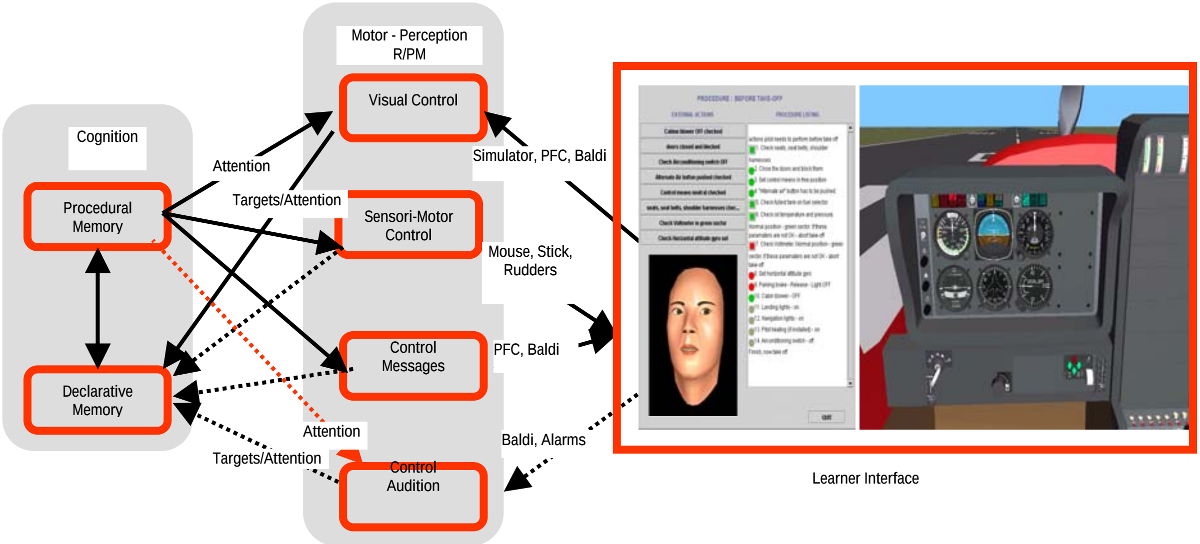

Figure 1. Cognitive control of dialogue's modalities by ACT-R/PM model in ASIMIL[1]

According to Piaget model, the declarative knowledge is posterior to the procedural knowledge. In the model ACT (Component of Though Activates) of John Anderson, (here we use ACT-R/PM of Anderson & Byrnes) [1], the formalised articulation is opposite to the processes of knowledge acquisition. The cognition is analysed on the symbolic level and the basic cognitive unit is the condition-action rule.

The working memory of R/PM engine is only seen as a system, which is independent of the long-term working memory. Cognition is then considered as a succession of cognitive processes which are connected dynamically, posing with a great acuity the problem of centralised control (Amygdala) or not (sensory effectors). The basic level is the activation of a concept, which characterises the state of a concept at rest. That level is more significant for experts than for non-experts.

The expertise of knowledge acquisition can be described like the sequential application of independent rules, which are compiled and reinforced by the exercise of automation, thus allowing the acquisition of procedural knowledge. Moreover, in ASIMIL, we have needed to control several methods of parallel dialogue and exchanges (messages – texts / word, orders – mouse / stick / caps / instructions, alarms – visual / sound...). Thanks to the model, one can also specify the role of the cognitive resources in the high level cognitive tasks and adopt proposals exchanged at the time of a conversation.

The ASIMIL trainee's environment must dynamically respond to ACT-R/PM's outputs and thus must also often be simulated at a high degree of fidelity. The knowledge that must be provided to ACT-R to complete a model of a person in an environment is essentially of two types: declarative and procedural. Declarative knowledge, is represented in symbolic structures known as *chunks*. Procedural knowledge, sometimes referred to as "know-how," is stored in symbolic structures known as *production rules*.

We have used the interaction in a manner of ACT-R/PM model, which provides an integrated model (module of cognition connected to perception-motor module) and a strong psychology theory on how interaction occurs. Furthermore, ACT-R/PM model allows to produce diagnostics in real-time, what is very important in the context of aeronautic training exercises, which are often time-critical.

Theoretical modelling of error

Often, methods of systems' design, applied to the modelling of the human operator, give the results too oriented towards the system, and not oriented towards the individual. Among the models used we can mention the following: scalar, overlay, error-based, genetic [7]. Usually, the activity of a human operator is not subject to the strict mathematical laws. It uses cognitive factors (motivation, stress, emotions), and cannot be evaluated efficiently via conventional mathematical equations. It becomes necessary to use other techniques resulting from research in the field of belief-based logic [8] or qualitative modelling [9].

Our study is centred on the analysis of trainee's behaviour as a base for modelling. The committed errors are used as reference marks for the detection of changes in behaviour. Three reference frames are defined to cover various solutions which the trainee is able to adopt in the course of the realisation of a given task. These reference frames can be used jointly [10]: (a) frame of the prescribed task: the error is defined like a non-respect or non-application of the procedure, (b) standard frame of the typical user: this frame represents the set of tasks carried out by a standard operator in the same profession, (c) frame of the operator himself/herself (individual activity).

Consequently, we need a system which does not analyse only the errors due to the lack of knowledge, but which also carries out also a complete diagnosis of the operator. The general outline of the process of the evaluation is presented on the Figure 2.

divergences are analysed by the system in order to determine the level of trainee's self-confidence [11].

Learner profile represents a structure which includes different data, as in our study. Information is organised in 3 levels:

- course (global performance, advancement in the curriculum)
- exercise (history of all exercises)
- action (all realised actions during the exercise are recorded).

According to [12], one of the main characteristics of an ITS, as well as for the human teacher, is to be qualified in the subject which ITS teaches. For ITS, the approach consists in equipping the system with capabilities to reason on any problem of the field (within the limits imposed by the syntax of the command language). A major consequence of this projection is reflected by the overall evolution of computer-based instruction systems, which evolved from the principle of codification of pre-set solutions towards the processes of resolution.

In qualitative representation, it is important that

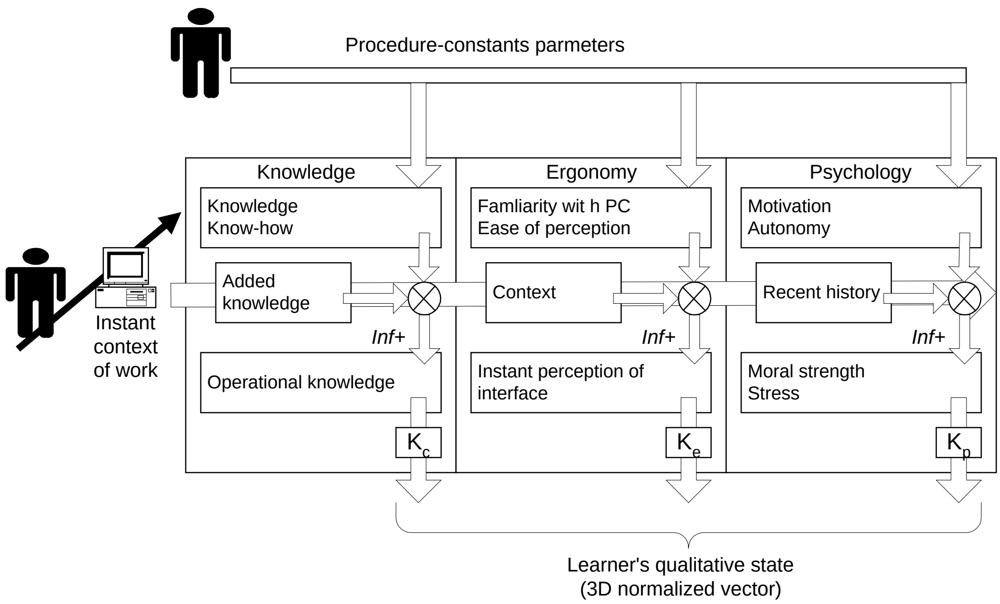

Figure 2. Process of qualitative evaluation

trainees become designers because during the process of design, they are brought to articulate the relations between the entities and the various beliefs about these entities. The suggested qualitative models must provide the means for beliefs' externalisation and the support for the reasoning, for the discussion and for the justification of decisions [9].

Another characteristic of our system of trainee's modelling consists in the fact that the evaluation and the integral note include the components of three kinds: knowledge, ergonomy, psychology (see Fig.2).

Inside each criterion, there are elements (for example, knowledge and know-how of the trainee in the criterion "knowledge") which are viewed as constant during training session.

A primary typology of gaps ("tunnel", intervals of motivation) was already established [13]. The analysis of works carried out by Rasmussen [3], Norman [14], Reason [15], brought us to extend this typology, by distinguishing three different types of errors / cognitive gaps:

- the errors due to an insufficient knowledge – this type of error represents the fundamental assumption of the result following an incorrect action/answer
- the errors due to a bad ergonomy of the ITS – these errors can be detected after the observations of the interactions of the trainee with the ITS
- errors in connection with factors psychophysiologic of the human operator (i.e. human factor). These

Then, at the time of training session, each criterion performs an analysis of the trainee's actions from its own "point of view" (that of the criterion in question).

The evaluation of each criterion gives a qualitative coefficient (Kc, Ke or Kp according to the name of the criterion – knowledge, ergonomy, psychology), which is used in the calculation of the current performance. The criteria are then integrated in a general criterion of evaluation, which value determines tutor's intervention.

According to the error's gravity, the graph of total performance is built online. Teacher's intervention is carried out according to score's and its derivatives' changes.

Moreover, the terms of surprised error and awaited error are introduced in order to be able to calculate the rate of error's expectation by the ITS. This coefficient is used in the process of decision-making – is this particular error expected by the ITS or not?

More coefficients K are high, more error's expectation is low (error's surprise is high). Thus, K determines the character of teacher's assistance provided to learner.

As learner evolves in three-dimensional space

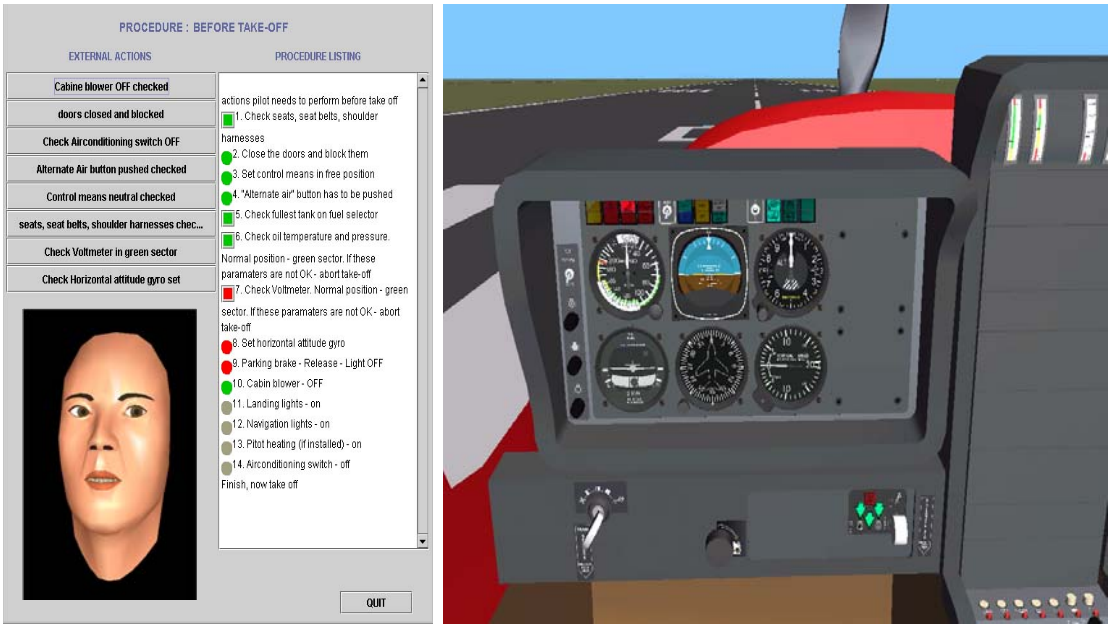

Figure 3. System of procedures follow-up on the left, flight simulator on the right and an animated agent (Baldi)

(knowledge, ergonomy, psychology), we have the possibility to follow his/her integral progression (by measuring instantaneous length of the vector of error *E* like its performances on each one of the criteria c, e or p, see also results presented in section 5).

Multi-Agent System Architecture

As ITS was designed under the form of a Multi-Agent System (MAS), this section briefly presents its main components.

Multi-agent technologies are widely used in the field of ITS [16 ; 13]. The aeronautical training has five characteristics which make it particularly adapted to agent applications: it is modular, decentralised, variable, ill-structured and complex. Three main components of an ITS (student, knowledge and pedagogical models) were integrated in the architecture of intelligent agents such as Actors [16]. This architecture is presented on the Figure 2, and the interface of the whole system on Figure 3.

A dedicated architecture for perception and qualification of errors has been constructed around the adaptation of the ACT-R/PM (Byrnes, Anderson 2001), which consists in the cognitive tracking based on intelligent agents.

The various agents of this architecture are:

- Interface agent ensures the communication between the MAS and the other components of the system (simulator, virtual reality, procedures)
- Curriculum agent traces the evolution of learner in interaction with the system and builds history
- team of agents-Evaluators of errors realises diagnoses of trainee's errors according to three axes: knowledge, ergonomy or psychology
- Pedagogical agent carries out the evaluation and brings a help to learning
- agent-Manager of the didactic resources looks up for pedagogical resources required.

The effectiveness of the follow-up by agents ASITS was already shown in CMOS prototype [13]. Today, the presence of several agents-evaluators allows the diagnosis of several types of errors. The agents-evaluators launch the evaluation, then the variation is quantified, evaluated and redirected towards the Pedagogical agent in order to be operated (announced and/or stored for the debriefing).

The system ASIMIL was the object of evaluations during 8 months in real conditions. These evaluations have involved trainees and private pilots from France and Italy (since they were conducted in the framework of

ASIMIL project). The evaluations allowed to underline the following tendencies:

- trainers perceive the tool positively: according to them, such a software could become of a good support for trainees (Pedagogical agent doesn't miss any error) and for trainers themselves (the agents' debriefing is explicit and can serve as a base for face-to-face debriefing)
- trainees also have approved the software, but they pointed out the disturbing character of Pedagogical Agent who spoke in English only. In reality, the training is often performed in native language (French or Italian, in our case) even if international aeronautic requirements (JAR, FAR)[1] are formal and recognise English as the only official training language.

Example of agents' functioning in aeronautic training

Figure 3 shows three components of training environment. The procedure presented here is the procedure of takeoff ("Takeoff Procedure"). The trainee must carry out a series of actions on the simulator, whereas the system of procedures follow-up PFC validates the actions carried out by learner. If learner's action does not correspond to the action required by the procedure, a red light is displayed.

The Pedagogical agent (animated character on Figure 3) carries out the teaching expertise on trainee. Its diagnoses are based on the trainee's history. The animated agent was developed in co-operation with the university of Santa Cruz, California.

In complement of traditional means of trainee's evaluation, we concentrated our efforts on the means at the disposal of the human tutor (instructor). Two main functions were identified as essential for an effective follow-up of a trainee by the trainer:

- synchronous or asynchronous follow-up of the trainee: show of the events recorded in the history of each training session (see Figure 4)
- customisation of Pedagogical agent: changing its physical appearance (face aspect, gravity of voice) as well as its reasoning (thresholds of help messages, conditions of stopping procedure).

Two undeniable advantages consist in the fact that the agents do not let pass any deviation/error, and in carrying out, for the instructor, the supervision of several trainees simultaneously.

The following details are presented in the window of instructor (see Fig.4). The axis of abscissa means time starting from the beginning of the exercise. The axis of ordinates means the variation of the objective of the exercise (also called user's "qualitative score"). A certain number of general options enters in account, such as level of learner, mode of training, tolerances, coefficients of performance Kc, Ke, Kp etc. The monitoring table (in the middle of each panel on Fig.4) holds the chronology of the session. One can see the moment when an error has appeared (column "Temps"), the qualification of the error ("Diagnostic"), its gravity ("Gravité", a number of points to be removed, associated with gravity with the error – slight, serious or critical), degree of error's expectation ("Attente"), and proposed help ("Aide").

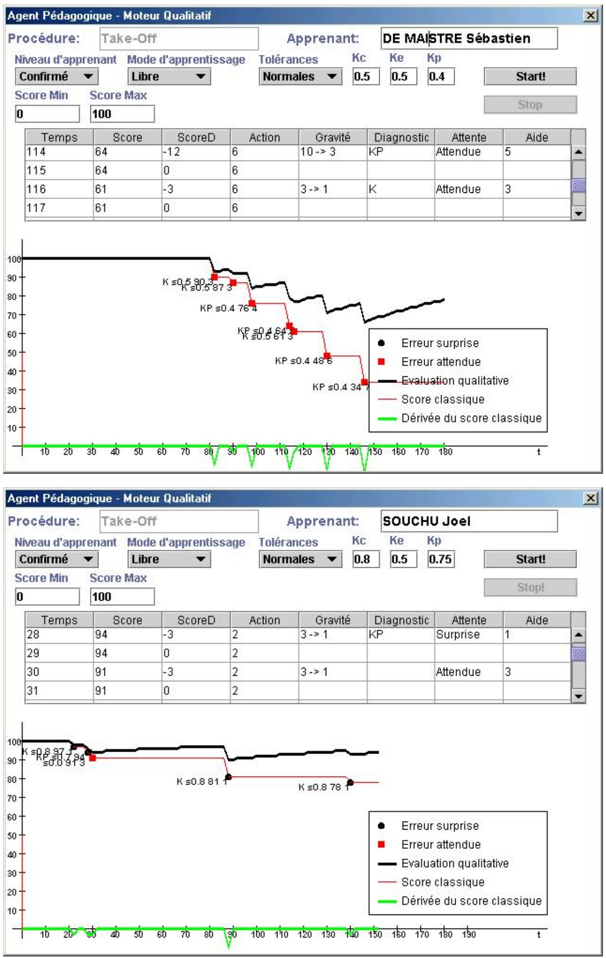

Figure 4. Instructor's « Dashboard » cases

The analysis of the curves shows that:

- on the panel above, the session is performed by a learner with the high level of knowledge ("confirmé"), but rather weak Kp, which seems to be confirmed by the error count of the type P (psychology). This trainee has been lost facing an error but, after some hesitations, has found the correct solution of the exercise
- on the panel below, the session is performed by a regular trainee, who made two errors, but quickly found the ways of correcting them.

The analysis of the curves of performance by the instructor not only makes it possible to evaluate learners, but also of re-adjusting the rating system of errors, by modifying weights of various errors. As an expected issue, the qualitative accreditation of differentiated users, can be done by reflexive comparison of the local deviation during the dynamic construction of the user profile. The analysis of the curves red and black allows

---

[1] The Joint Aviation Authorities (JAA) is an associated body of the European Civil Aviation Conference (ECAC) publishes the Joint Aviation Requirements (JARs) whereas the Federal Aviation Administration edits the Federal Aviation Regulations (FAR)

to match similar patterns (or not) to be detected (manually in the current version) and the green curve give alarms to start the qualitative accreditation process.

Conclusions and perspectives

In the aeronautics, the complexity of an aircraft's cockpit as well as the number and disposition of its instruments get the task of computer-assisted training (and, in particular, of distance training) very difficult. In the most recent systems of cabin simulators, the evaluation and certification are based on the fact that the instructor follows the trainee step-by-step what causes the prohibitive costs of classroom training.

We presented an original step to provide an evaluation complementary to that which is traditionally obtained by the instructor. This step, like the majority of the other approaches, is possible thanks to the techniques resulting from various fields (ITS, MAS, AI, Modeling) but with the concern of keeping the reasoning close to human logic.

Characteristics like the follow-up in real-time by the instructor of the learner's cognitive discrepancies, by making the distinction between the errors relative to the simulator, to the procedures or to the cognitive tasks, established the base of our study. We wished and carried out a rational and rigorous taking into account of the variables and general options from the method of qualitative simulation, which led to the establishment of a hierarchy of errors representing a significant progress compared to preceding work. The extension of the errors' evaluation represents an unavoidable phase in the process of certification of the tool.

References

[1] M.D. Byrnes, J.R. Anderson, Serial modules in parallel: The psychological refractory period and perfect time-sharing, Psychological Review, 108, 847–869. 2001.

[2] J. Self. The Role of Student Models in Learning Environments. AAI/AI-ED Technical Report No.94. In Transactions of the Institute of Electronics, Information and Communication Engineers, E77-D(1), 3-8, 1994.

[3] J. Rasmussen. « Information processing and human-machine interaction : an approach to cognitive engineering », North-Holland, 1986.

[4] A.A. Krassovski. Bases of simulators' theory in aviation. Moscow, Machinostroenie, 1995 (in Russian), 304p.

[5] O. Popov, R. Lalanne, G. Gouardères, A. Minko, A. Tretyakov. Some Tasks of Intelligent Tutoring Systems Design for Civil Aviation Pilots. Advanced Computer Systems. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers. Boston/Dordrecht/London, 2002.

[6] F.E. Ritter, D. Van Rooy, F. St Amant. A user modeling design tool based on a cognitive architecture for comparing interfaces. Proceedings of the Fourth International Conference on Computer-Aided Design of User Interfaces (CADUI), 2002.

[7] P. Brusilovsky. Intelligent tutor, environment, and manual for introductory programming. Educational and Training Technology International 29 : pp.26-34, 1992.

[8] K. Van Lehn, S. Ohlsson, R. Nason. Application of Simulated Students : an exploration. Journal of Artificial Intelligence in Education, vol.5, n.2, 1994 ; p.135-175.

[9] P. Salles, B. Bredeweg. « A case study of collaborative modelling : building qualitative models in ecology ». ITS-2002, Workshop on Model-Based Educational Systems and Qualitative Reasoning, San-Sebastian, Spain, June 2002.

[10] F. Jambon. "Erreurs et interruptions du point de vue de l'ingénierie de l'interaction homme-machine". Thèse de doctorat de l'Université Joseph Fourier (Grenoble 1). Soutenue le 05 décembre 1996.

[11] E. Aïmeur, C. Frasson. Reference Model For Evaluating Intelligent Tutoring Systems. Université de Montréal, TICE 2000 Troyes – Technologie de l'Information et de la Communication dans les Enseignements d'ingénieurs et dans l'industrie.

[12] P. Mendelsohn, P. Dillenbourg. Le développement de l'enseignement intelligemment assisté par ordinateur. Symposium Intelligence Naturelle et Intelligence Artificielle, Rome, 23-25 septembre 1991.

[13] G. Gouardères, A. Minko, L. Richard. « Simulation et Systèmes Multi-Agents pour la formation professionnelle dans le domaine aéronautique », Dans Simulation et formation professionelle dans l'industrie, Coordonnateurs M. Joab et G. Gouardères, Hermès Science, Vol.6, No.1, pp.143-188, 1999.

[14] K.L. Norman. « The psychology of menu selection : designing cognitive control at the human/computer interface », Ables Publishing, Norwood NJ, 1991.

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

[16] C. Frasson, T. Mengelle, E. Aïmeur, G. Gouardères. « An actor-based architecture for intelligent tutroing systems », Intl Conference on ITS, Montréal–96.