COSIMO: A Cognitive Simulation Model of Human Decision Making and Behavior in Accident Management of Complex Plants

Pietro Carlo Cacciabue, Françoise Decortis, Bartolome Drozdowicz, Michel Masson, and Jean-Pierre Nordvik

***Abstract*—The principal way to improve the human–machine system is to work at the level of human contribution. The use of a cognitive model of human behavior for sustaining the design or the study of the interaction with the system under control is increasing. A cognitive simulation model (COSIMO) is described, which simulates the behavior of an operator controlling a complex system during the management of accidents. Particular attention is paid on the theoretical foundations of the model, on its computational implementation and on a number of simulations cases of man–machine system interactions. The approach followed in COSIMO is to build a structure for the various kinds of cognitive functions that are performed by an operator in complex environments, e.g., information seeking, pattern recognition and diagnosing, monitoring, planning and acting on the system. For the implementation of COSIMO a blackboard architecture has been chosen. This architecture offers enough flexibility and modularity for modeling the overall state of the human and of the environment.**

NOMENCLATURE

AFWS	Auxiliary feedwater system
A_j	The jth row of the EAM.
BB	Blackboard.
CBB	Control blackboard.
CF	Cognitive function.
CIF	Currently instanciated frame.
COSIMO	Cognitive simulation model.
$D(A_j, v)$	Diagnosticity of value v of attribute A_j.
DBB	Domain blackboard.
DI	Direct inference.
EAM	Entity-attribute matrix.
FG	Frequency gambling.
KB	Knowledge base.
KBF	Knowledge-base frame.
N	Total number of entities in the EAM.
n	Number of entities where the value A_j is v.
RBF	Rule-base frame.
SM	Similarity matching.
v	Value of an attribute.
WM	Working memory.

Manuscript received July 1, 1991; revised February 2, 1992.

The authors are with the Institute for Systems Engineering and Informatics, Commission of the European Communities Joint Research Centre, Ispra-VA 21020 Italy.

IEEE Log Number 9201412.

I. INTRODUCTION

THE INTERACTION between operators and the plants they control has nowadays become a central issue in design and safety analysis, even if the improvements of technology contribute more and more to the development of automatic systems that assign to the operators the sole role of supervisors [1]–[4]. This trend may be justified by the consideration that the vast majority of the accidents is the result of a combination of factors, inevitably involving the contribution of human erroneous and/or inappropriate behavior. In many cases, such as when the human erroneous actions are combined with inappropriate or limited protection and safety measures, the consequences of accidents are impressive for the environment and for the extent of casualties. On the other hand, the principal way to improve the reliability of the human–machine system is to work at the level of human contribution rather than at the machine component [5]–[7]. This problem has been tackled from different perspectives:

1) in control and plant management, by designing and developing decision support systems;
2) in applied ergonomics, by designing control rooms and advanced human–machine interaction tools;
3) in reliability analysis, by studying safety of plants and procedures including the crucial effects of human erroneous behavior.

The common feature of these three main methodological approaches lies in the use of a model of human behavior for sustaining the design or the study of the interaction with the plant under control. There exist several models related to the human operator in a work environment that have followed different trends. Some models are strongly affected by information technology without paying much attention to data coming from field studies. As an example, CES [8], based on Caduceus diagnostic system, and OFMspert [9], implemented in a blackboard architecture (BB), are oriented toward decision support systems. They are sustained by cognitive theories and they interact with a simulated dynamic environment. Other models take as a starting point the data collected in real situations and formalize a model on this basis. AIDE [10], a model of a pilot aircraft, is an example of such type of approaches. Finally there are models, exclusively field oriented, which aim at guiding the design of work environment. As an example, the model of Pinsky and Theureau [11] excludes any implementation of the formalism.

The approach followed in the cognitive simulation model (COSIMO), currently developed at the Ispra-Joint Research Centre of the Commission of the European Communities, is to identify and model the cognitive functions that are performed by an operator in complex work environments during accident management, e.g., information seeking, pattern recognition and diagnosing, monitoring, planning and acting on the system. While many existing modelings [12]–[14] focus separately on models of detection, planning, diagnosis or execution with different formalisms, in COSIMO these cognitive functions are integrated and become operative in a modeling environment according to some primitives of cognition that are assumed to represent the "mechanisms of conservation" of the cognitive processes [15], [16].

These cognitive functions have to be instantiated in a rich informatic structure offering enough flexibility and modularity to model the overall state of the human and of the work environment. To this aim, the blackboard architecture [17], [18] has been chosen to describe and enable the interaction of various sources of knowledge and mechanisms that are required for the simulation of the decision making process [19]. COSIMO runs on a lisp machine and it interacts with the simulation of the plant under management and control, which is implemented on a Sun computer.

The basic application of COSIMO is to explore human behavior in simulated accident situations in order to identify important safety aspects of the human–machine interaction system. In a complete and powerful modeling architecture of the pant and operator behavior, COSIMO could be usefully adopted to:

- analyze how operators are likely to act given a particular context,
- identify situations that can lead to human errors and evaluate their consequences,
- identify difficult problem solving situations, given problem solving resources and constraints (operator knowledge, man–machine interfaces, procedures),
- identify and test conditions for error recovery,
- investigate the effects of changes in the man–machine system.

The objective of this paper is the presentation of the state of the art of COSIMO, describing firstly its theoretical foundations and computational implementation. Then, the current simulation possibilities and limits of application of COSIMO are discussed in a number of sample cases of man–machine system interactions. In these cases the physical system under control is a typical nuclear power plant subsystem, the auxiliary feed-water system (AFWS), which has been reported somewhere else [20] for the model of the physical phenomena and for the simulation of the control instrumentations and procedures in accident management.

II. Theoretical Foundation of COSIMO

A. Architecture of Cognition

1) Working Memory and Knowledge Base: The cognitive model of the operator assumes an information processing architecture that consists of two parts (Fig. 1) [15]: the working memory (WM) and the knowledge base (KB). The KB can be seen as a virtually limitless repository of information containing both declarative and procedural knowledge structures. These structures capture the experience gained by the operator during both his day to day work and his formal and practical studies. The WM is a limited, serial, working area associated with powerful computational facilities for the management and temporary storage of data required by the cognitive processes. By a cognitive process we intend any operator's behavior and decision making activity carried out in the complex work environment. The role of the WM is: to analyze and assess information about the environment with respect to the content of the KB; to generate hypotheses; to decide whether these hypotheses are appropriate; and if need be to repeat the search with revised assumptions.

The cognitive model of the operator assumes an information processing architecture that consists of two parts (Fig.1) [15]: the working memory (WM) and the knowledge base (KB). The KB can be seen as a virtually limitless repository of information containing both declarative and procedural knowledge structures. These structures capture the experience gained by the operator during both his day-to-day work and his formal and practical studies. The WM is a limited, serial, working area associated with powerful computational facilities for the management and temporary storage of data required by the *cognitive processes*. By a cognitive process we intend any operator's behavior and decision making activity carried out in the complex work environment. The role of the WM is to: analyze and assess information about the environment with respect to the content of the KB; to decide whether these hypotheses are appropriate; and, if necessary, to repeat the search with revised assumptions.

2) Rule-Based Frames and Knowledge-Based Frames: The basic components of the knowledge base are rule-based frames (RBF) and knowledge-based frames (KBF).

The RBF's were originally conceived as static frames (i.e., snapshots of the configuration of the process controlled by the operator) for diagnosing and recovering from an accident [16]. Each RBF consists of the following elements (Fig. 2):

- an index label identifying the type of accident (e.g., loss of electrical power, steam generator tube rupture);
- a subjective frequency tag related to the number of times the operator has encountered this accident in the past as well as his familiarity with the accident;
- a set of properties or attributes, and their associated attribute-values describing the symptoms expected and characterizing such an accident (e.g., pressure of water in the steam generator : high and increasing);
- a set of actions appropriate to the situation.

Two types of symptoms can be described in a RBF: physical symptoms and logical symptoms. Physical symptoms

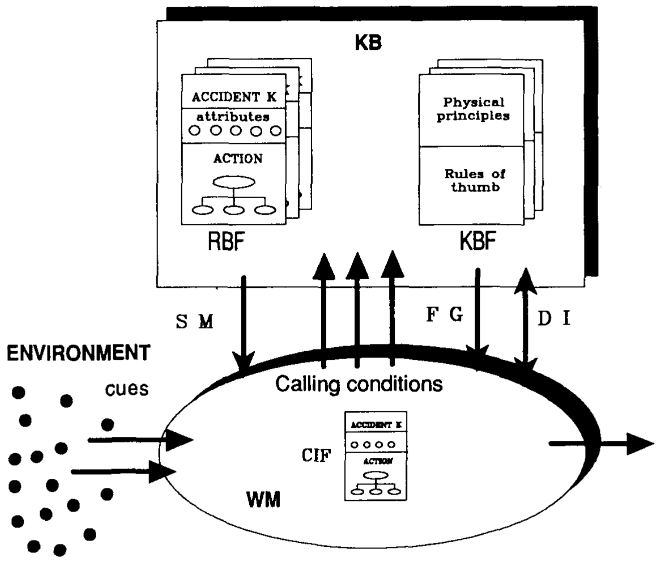


Fig. 1. COSIMO structural description.

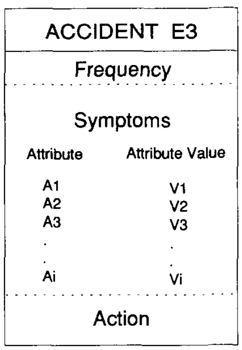


Fig. 2. A rule-based frame.

are related to values of plant parameters (e.g., mass flow to steam-generators is high, pressure in steam-generator has decreased) and logical symptoms to Boolean indicators (e.g., pump indicator is on, regulation valve indicator is off). A symptom consists of the following elements:

- variable (e.g., pressure),
- material (e.g., water),
- plant component (e.g., SG1),
- numerical value (e.g., 55),
- semantic value (e.g., high),
- numerical derivative (e.g., +4 u/t),
- semantic derivative (e.g., increased),
- time (s),
- numerical value history (e.g., 200, 300, 350),
- physical salience,
- cognitive salience

Physical salience reflects the physical properties of a symptom (e.g., type of indicator, position in the control panel, light or sound intensity of alarms).

Cognitive salience represents the subjective familiarity and significance of a symptom. For instance, a particular indicator may be checked by the operator because he knows that this indicator is of particular relevance for the current task.

In order to represent the dynamic evolution of the process a further knowledge structure is defined—a dynamic frame—as a time-dependent extension of a static RBF. A sequence of these dynamic frames describes the evolution of a particular process configuration, i.e., the various steps of the cognitive process carried out by the operator while controlling the plant. In addition to the elements of the static frames the dynamic frames comprise also (Fig. 3):

- a time stamp characterizing the instant at which a snapshot of the accidental situation is taken: this is the time elapsed since the beginning of the event;
- persistent symptoms observed over time—unchanged symptoms;
- newly appearing symptoms, not previously observed—new symptoms;
- symptoms that were already present but have changed over time—changed symptoms.

KBF's are packets of knowledge containing only heuristic rules of thumb as well as general engineering and physical principles on the operation of the plant, usually developed during training, experience and theoretical background.

Contrary to the RBF's, KBF's do not contain organized sets of recovery actions and symptoms associated with well-defined situations, but they are called into play in the WM when a planning process has to be developed either by adapting to a novel situation known parts of plans or knowledge that worked correctly in analogous circumstances or by developing a totally new strategy of actions.

This part of the modeling architecture represents one of the most difficult parts of theoretical development of the model

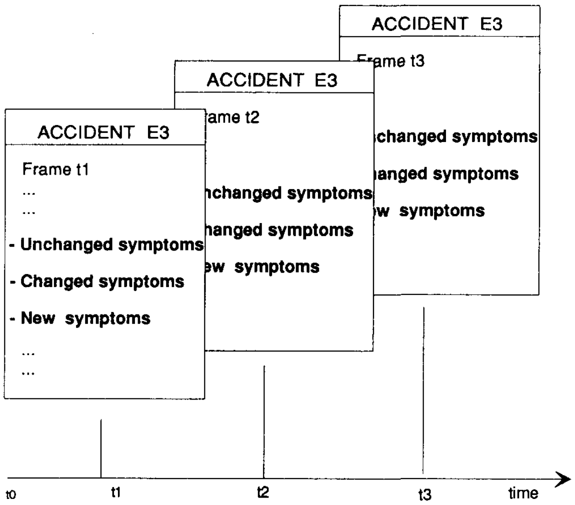


Fig. 3. Dynamic rule-based frames.

Attribute \ Entity	E1	E2	E3	E4	.....	Ej	...	EN
A1	V1	V1	V1	V1	...	V1	...	V1
A2	V1	V4	V3	V5	...	V1	...	V9
A3	V4	V2	V3	V7	...	V5	...	V8
....	...	...	...	...	...	...	...	...
Ai	V1	V3	V5	V1	...	V6	...	V5
...	...	...	...	...	...	...	...	...
Am	V6	V3	V2	V5	...	V2	...	V2
Frequency	0.8	0.8	0.7	0.6	.....	0.5	...	0.8

Fig. 4. The EAM.

and it is still under detailed study and revision before the actual formalization in logical and computational expressions. Consequently, it will only marginally be discussed in the remaining part of this work.

3) Entity Attribute Matrix: A synthetic view of the knowledge about diagnosis contained in the KB can be constructed over the collections of RBF's. This structure—the entity-attribute matrix (EAM)- takes the form of a two-dimensional matrix, with entities as row heads and attributes as column heads [21]. Each entity is associated with a RBF and is characterized by the set of attributes/attribute-values, i.e., symptoms, and the frequency tag of that RBF. An element of the matrix represents therefore the value of a particular attribute in a particular accident situation. An example of EAM is given in Fig. 4.

Since the set of attributes describing each entity differs from one entity to another, empty spaces are left in the EAM concerning value of attributes not present as attribute of the corresponding entity. The number of attributes, and thus the number of empty spaces, of an entity is related to the knowledge that the operator has about the accident associated with that entity. This knowledge depends on the number of times the accident has been encounter by the operator: the higher the accident frequency, the more numerous the descriptive attributes. This refers to Reason's experimental finding "the more frequent in the world, the more will be known about it" [22].

Each attribute of the EAM is characterized by a domain of variation, i.e., the set of different values taken by the attributes over the entities. The cardinality of these domains varies from one attribute to another. A particular value v of an attribute A_j is characterized by its diagnosticity $D(A_j, v)$ which is an easy-to-compute information measure given by

$$D(A_j, v) = (N - n)/(N - 1)$$

where N is the total number of entities and $n(0 < n \leq N)$ is the number of entities where the value of A_j is v. The diagnosticity of an attribute value reflects the level of discrimination provided by this value over the entities. The more a value is present, the less "informative" is the value and the smaller is its Diagnosticity. In particular, the limit conditions are: $n = N \Rightarrow D = 0$ and $n = 1 \Rightarrow D = 1$).

This simple equation has the objective to facilitate the understanding of the concept of pertinence of information, as it has been used in this work. A more formally complete and mathematically elegant framework for developing the notion of diagnosticity may be given, for example, by the information theory, which provides an appropriate metric in this case [23], [24].

B. Primitives of Cognition

The basic assumption of COSIMO concerns the Cognitive Primitives or mechanisms by which the frames of the KB are selected and brought into the WM for further processing. Two primitives have been identified so far, namely the similarity matching (SM) primitive and the frequency gambling (FG) primitive, which are closely related to the underspecification theory of Reason [15].

The SM primitive matches information—cues—coming from the work environment with diagnostic attributes described in the KB. More precisely, the SM compares external cues, i.e., data usually perceived because of their high physical salience, and internal cues, i.e., data usually searched for because of cognitive process, with previously experienced symptoms contained in KB. The result of the matching is either a unique hypothesis or a set of partially matched hypotheses, depending on the discriminating ability of the considered cues. These hypotheses are characterized by a matching value that represents the result of the SM primitive

The FG primitive resolves the conflict, which may occur between partially matched hypotheses selected by the SM, in favor of the most frequently encountered and well known accidental situation.

These two primitives operate so as to minimize cognitive strain and maximize the chances of immediate information processing, which is often recognized as the most common problem solving strategy of operators in emergency situations. The problem solving based on the SM-FG primitives is therefore defined as an immediate information processing by contrast with a deep reasoning process. Switching from immediate information processing to deep reasoning is assumed to be possible only when sufficient time and attention resources are available. In the test cases addressed so far, only immediate information processing have been considered. As a matter of fact in critical situations the operator has to find an appropriate solution in a limited time and therefore may be extremely urged to make a decision. It is therefore assumed that in such a situation no sophisticated reasoning takes place and that the operator resorts only to previous experience on similar cases.

However, not all human problem solving activities resort to pattern recognition only and therefore a third cognitive primitive must also be considered—the direct inference (DI) primitive—whereby decision and planning processes based on analogical and inductive/deductive reasonings are described. This primitive is activated only if the cognitive and environmental constrains allow the performance of such processes and when the result of a SM-FG primitives failed to identify an already compiled plan of actions. In these cases, the problem solver has to elaborate a strategy using a mostly serial process of inference, which requires a lot of attention resources and has been proven to be slow, tedious, and arduous.

In the cognitive architecture of COSIMO, there exists a strong modeling connection between the DI primitive and KBF's. The DI primitive represents the driving mechanism by which the knowledge contained in the KBF's is called into play in the WM during reasoning and planning processes, where the use of rules of thumb and of physical and engineering principles are used to shape analogical reasoning, temporal reasoning, qualitative reasoning and homomorphism representations [25], [26].

C. Functions of Cognition

Field studies in different industrial plants have shown that operator behavior consists generally of different Cognitive Functions (CF's), like diagnosing, monitoring, scheduling, etc. [27]–[29]. These functions are partially interdependent. For example, operators' perception is calibrated by their current view of the world and the way the problem in evolution is perceived. But in turn, as perception is selective, it tends to shape and sustain the current hypothesis against contradicting cues.

In the present version of COSIMO, four interrelated CF's have been described, namely Filtering, Diagnosing, Hypothesis Evaluation and Execution. These functions rely on the one hand on information of what is known about operator's strategies in real situations [30]–[33], and on the other hand on what has been experimentally investigated in cognitive psychology [34].

1) The Filtering Function: The communication between the plant and COSIMO is done through a "cognitive filter". The filter selects among the large number of data produced by the environment those which are actually perceived. This perception is guided by a salience criterion based on the physical and cognitive salience associated with the environmental data [35]. After the filtering process, data are interpreted: numerical data are translated in semantic information (e.g., nominal, low, on, off, etc.). Only filtered data are interpreted according to the psychological argument that only the information "perceived" is actually interpreted.

2) The Diagnosis Function: Once the cues have been filtered and interpreted, they are matched with the KB using the SM mechanism in order to identifying a RBF to be executed. The number, form and quality of these cues determine the matching value: the more pertinent the calling conditions, the less numerous and the better supported are the hypotheses selected from the KB. Once a frame is activated in the WM, it will govern a hypothesis driven information processing, by which a confirmation process is initiated according to the attributes that better qualify the selected frame. In case of ambiguity, either in the interpreted data or in the symptoms represented in the KB, more than one explicating hypothesis may be brought to mind by the SM. The FG resolves the conflict between partially matched hypotheses in favor of the more frequent one encountered in the past.

Matching and frequency values are combined to deliver a support score for each hypothesis. This combination may take various forms according to the level of expertise and current cognitive state of the operator, i.e., the mental and physical conditions characterizing the operator as a consequence of his interaction with the working environment and the sociotechnical context. For instance, an expert operator may always favor the frequency value whereas a novice may give higher priority to the similarities between situations. The hypothesis that is eventually selected is the one obtaining the largest support score.

The diagnosis is intended to be a dynamic process that can be carried out during an extended period of time, interrupted for some alternative plan or new diagnosis and possibly resumed at a later stage. A diagnosis can also be triggered at any time during the execution of another function just to confirm the initial choice of a RBF.

3) The Hypothesis Evaluation Function: This function aims to decide whether an hypothesis can be trusted or has to be rejected. If the support of the hypothesis selected after the diagnosis function is too low to provide the model with enough confidence in this explanation, the hypothesis is rejected and a new diagnosis is initiated between the already perceived cues from the environment and the known cues from the KB. The level of confidence required for an hypothesis to be considered—evaluation threshold—can be dynamically modified by the model according to the current operator's cognitive state.

4) The Execution Function: Once an hypothesis has been selected, the WM is cleaned out and receives an instantiation of the RBF associated with the selected explanation. Therefore only one unique RBF can be present in the WM at a time. This RBF is called the currently instantiated frame (CIF). The recovery actions compiled in the CIF are then executed over the time evolution of the accident. Note that a recovery action does not necessarily implies a physical action upon the plant (e.g., a monitoring process).

D. Cognitive Interaction

Several researches have shown the CF's are not necessarily sequentially concatenated. On the contrary, depending on their level of expertise, operators are able to make short-cuts [14], [34]–[36]. In particular, operators may work forward and backward: they may focus on a first diagnosis, perform some actions and, when unsatisfied with the resulting effects, modify their initial diagnosis. Sometimes a "dilatory action" is performed to delay the evolution of a sequence when the operator wants to extend the available time for understanding the current situation. As a consequence CF's cannot be strictly ordered along the temporal axis but should intervene in an opportunistic way.

Such an interaction between CF's is guided by higher level cognitive functions that are contextual to the specific work environment of application of the model [37], [38]. In particular, such functions have to take into consideration the modes of control and the specific competencies that are developed by the operators in different working environment and become consequently a consistent part of the "culture" of those individuals. These special functions have to be accounted for in the implementation of the model as the controlling component of the simulation.

However, if the evolution of the accident follows the expectation of the operator and thus the information contained in the CIF, the interaction is substantially sequential and follows the order filtering, diagnosing, hypothesis evaluation and execution at predefined time instants.

III. Computer Implementation

A. The Blackboard Paradigm

A blackboard framework is proposed as computational architecture to implement the cognitive model [17], [18]. This architecture has been chosen because of the following characteristics: 1) it supports the adaptive behavior of reasoning of an operator; 2) it supports an explicit representation of mechanisms that enables the dynamic revision of its behavior according to new information that continuously change the problem scenario; 3) it allows an incremental and opportunistic decision making development.

The blackboard (BB) architecture is often described as being composed of three parts, [39], [18] namely the blackboard itself with its internal structure; the Agents, also called knowledge sources or specialists, working on BB and the Controller governing the actions of the agents.

1) The Blackboard Structure: The blackboard structure consists of different levels of information that correspond to the various kind of structured-objects that may be dealt with by the model. These levels create a hierarchy along which the reasoning process takes place while the properties of the data describe the level at which objects are intended to be placed. Thus, the BB constitutes a global unique database seen by all the agents of the system. The set of all the properties is the restricted, well-defined vocabulary by means of which the model expresses its various actions and results (Fig. 5)

2) Agents: The overall inferential know-how of the domain is split into small independent units, each of which is specialized in a particular inference scheme. Each agent represents one of these domain-dependant inference mechanisms and focuses only on the BB levels of interest for its particular specialization. The agents are kept independent of each other and communicate indirectly through the BB.

An agent is mainly composed of three parts: the triggering body, the precondition body and the action body. In the triggering body, the initial conditions by which the agent can help the reasoning process are evaluated. An agent, even if triggered by some event, is not always able to directly perform its action body. Indeed, supplementary data may have to be acquired in order to fulfil the conditions of execution. These required conditions form the precondition body of the agent. The action body contains the knowledge on how to perform the inference mechanism on the data that match its triggering and precondition parts. The only way to modify the BB is during the execution of the action phase of an agent, either by creating new objects on the BB or by modifying previously existing objects.

Each time a new incoming data of the BB matches the triggering part of an agent, this agent generates a task. Such task contains a set of data that are used during the schedule process as well as during the possible execution of the action body of the agent. This task informs the Controller of the potential, but not yet ready, inference of the agent.

Thus, the execution of an agent takes place in three steps:

1) the agent is triggered and generates a new-task, called "triggered task"
2) the triggered task remains asleep until its precondition part is satisfied—these tasks are called "invocable tasks"
3) if the invocable task is selected using the conditions and criteria generated by the controller, the action part of the agent is executed.

Fig. 6 represents the basic BB cycle that takes place each time the Controller has selected a new task to execute.

The basic BB cycle consist of the following steps: 1) a particular task, called the current task of the BB-cycle, is selected by the Controller among the set of invocable tasks; 2) the action body of the agent associated with the current task is executed in the same environment that has matched its triggering and precondition bodies; 3) new-objects, or modifications of objects, are generated and update the BB; 4) new-tasks are generated by the agents that are triggered by the new incoming objects; 5) an updated set of invocable tasks is build by incorporating the new-tasks into the previous set of invocable tasks.

The dynamic behavior of the BB system can be understood as the chaining of such BB cycles, one after the other. The tasks generated along the course of action of the BB are placed into an agenda that is updated at each cycle. Inside the agenda, a distinction is made between triggered tasks and invocable tasks, which are the only tasks taken into consideration during the final stage of the selection process.

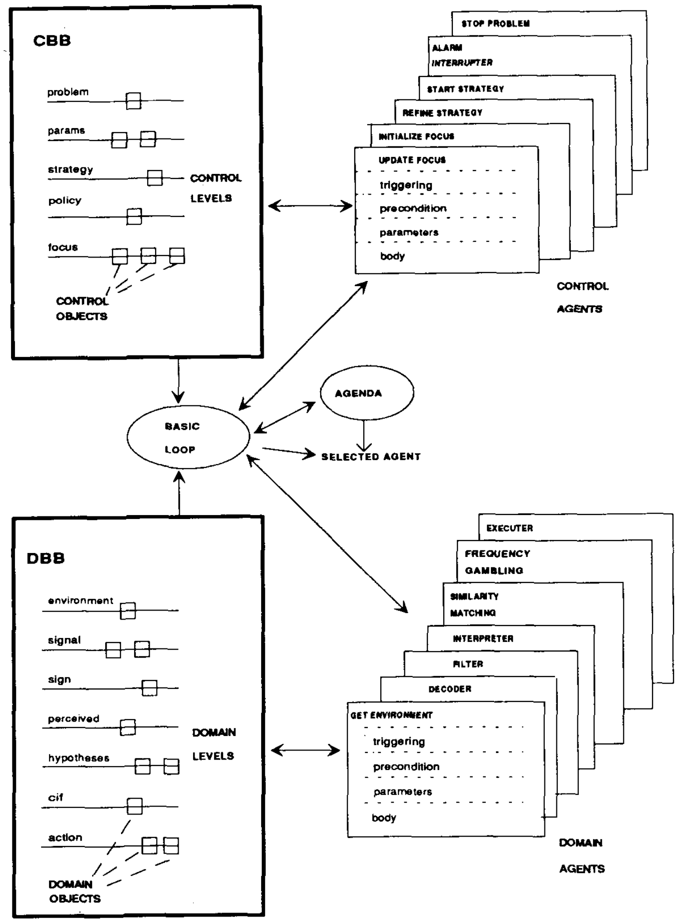


Fig. 5. The architecture of COSIMO.

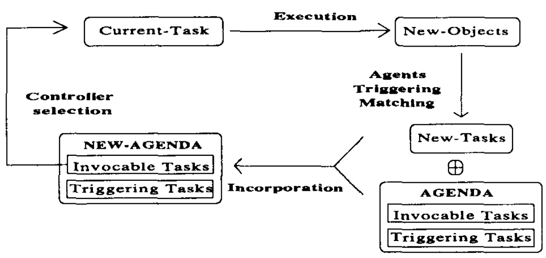


Fig. 6. Blackboard basic cycle.

3) Controller: As more than one agent can be triggered by the same object of the BB and as more than one object can be generated by a single task, a control mechanism is required in order to manage their interaction. The control mechanism is able to select amongst various invocable tasks the one to be activated next. It is obvious that this mechanism is of major importance in a BB system. Indeed the tasks are due to agents dealing at different levels of information and, therefore, have completely different meanings. Since the results, as well as the execution of the tasks are highly different, the product of the problem-solving process is directly related to the choice of the Controller: the more pertinent the selected task the faster the plant will converge toward an accurate behavior.

As explained by Hayes-Roth [17], the controller itself can be designed using a BB architecture. In this case, the control knowledge of the model and the different control levels, at which the control-process takes place, are explicitly represented by a BB structure and a set of agents. Different strategies and sets of heuristics refining these strategies can be applied to guide the decision process. The behavior of the model is built up in an incremental and opportunistic way as the BB and the Controller evolve. Therefore, in the same BB structure two different BB's coexist and interact with each other: a domain-BB (DBB) and a control-BB (CBB). In particular, the CBB generates tasks that are added to the tasks of the DBB, and then, given that the CBB governs the scheduling mechanism by selecting the task to be executed next, it may occur that it will choose one of its own task in alternative to others. In this way the CBB may influence its own scheduling mechanism. The various control agents and control levels of the CBB are briefly presented in Tables I and II.

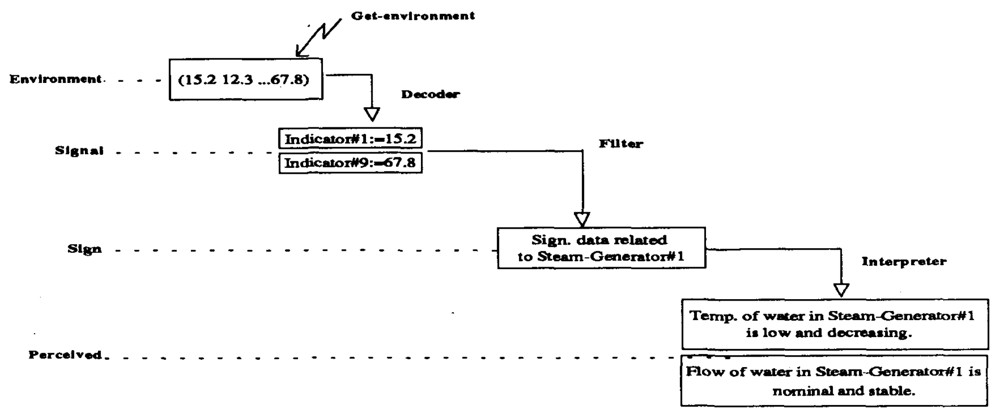


Fig. 7. Example of the data acquisition activity.

TABLE I
DESCRIPTION OF CBB LEVELS

Levels	Definition	Example
Problem	Problem to be solved	Maintain the safety levels
Strategy	Strategies currently used	Data acquisition,diagnosis, confirmation, . . .
Focus	Temporal agent selection criteria	Prefer tasks at signal level
Policy	Permanent agent selection criteria	Prefer recent tasks versus old tasks
Agenda	Set of invocable and triggered tasks	Get-Environment, Decoder, Update-focus, . . .
Task	Ccurrent task to be executed	Update-focus

TABLE II
DESCRIPTION OF CBB AGENTS

Agents	Triggering	Preconditions	Body
Put-problem-on-BB	None	None	Initialize BB activities
Stop-problem	New problem	Solution of triggering problem	Stop BB activity on current problem
Start-strategy	New problem	None	Activation of initial strategy
Update-strategy	New strategy	Rules for activation of new strategy	Selection of a strategy
Initialize-focus	New strategy	None	Selection of the first focus of triggered strategy
Update-focus	New focus object	Satisfaction of triggered focus	Selection of next focus

For example, "Update-Strategy" decides which strategy to adopt on the basis of the current configuration of the BB. Once a strategy has been selected, "Update-Focus" generates and updates, on the focus level, the various criteria that will sequentially implement this strategy.

This framework enables the system to tackle dynamic situations, where changes of lines of reasoning and changes of control strategies are needed to cope with the intrinsic nonmonotonicity of the problem. In the case of COSIMO this means that no sequence of cognitive primitives or functions are defined *a priori.*

TABLE III
DESCRIPTION OF DBB LEVELS

Levels	Definition	Example
Environment	Untreated representation of physical world	300, 9.7, 11, . . . 10.8, 0.1 . . .
Signal	Data decoded, set of indicators and alarms	– indicator 5=10.8 –alarm 23
Sign	Data cognitively and physically filtered	–indicator 5 – alarm 23
Perceived	Significant data interpreted (perceived cues)	Perceived problem relative to steam generator 1
Set of Hypotheses	Potential hypotheses related to perceived cues	–SG1 inlet break –SG1 inlet break and valve 1 blocked open
Current Hypothesis (Cif)	Hypothesis currently being used	SG1 inlet break
Action	Actions to be executed on the plant	–close valve v031 – open valve v035 – regulate SG1 at 30

B. Implementation of COSIMO

1) Architecture: The various agents and levels of the DBB are described in Tables III and IV. Currently, four correlated scenarios are simulated, namely: data acquisition, diagnosis, confirmation and planning-execution.

These strategies are defined by means of their related levels and agents at the DBB level and implement the SM and FG primitives as well as the four cognitive functions. The DI primitive has net been implemented so far. According to the BB formalism, each strategy is represented by an object at the strategy level. In addition the different stages of each strategy are focus objects of the focus level. The dependence amongst these scenarios, i.e., their interaction (interruption, chaining, shifting), are governed by the CBB. So far only very simple mechanisms of control have been implemented.

The data acquisition scenario: This scenario implements mainly the filtering cognitive function and consists of the following stages (Fig. 7).

The uninterpreted input data from the environment (i.e., external cues), are represented by objects of the Environment Level. This task is performed by the Get-Environment Agent (see Table IV). The Decoding Agent relates these raw data (a

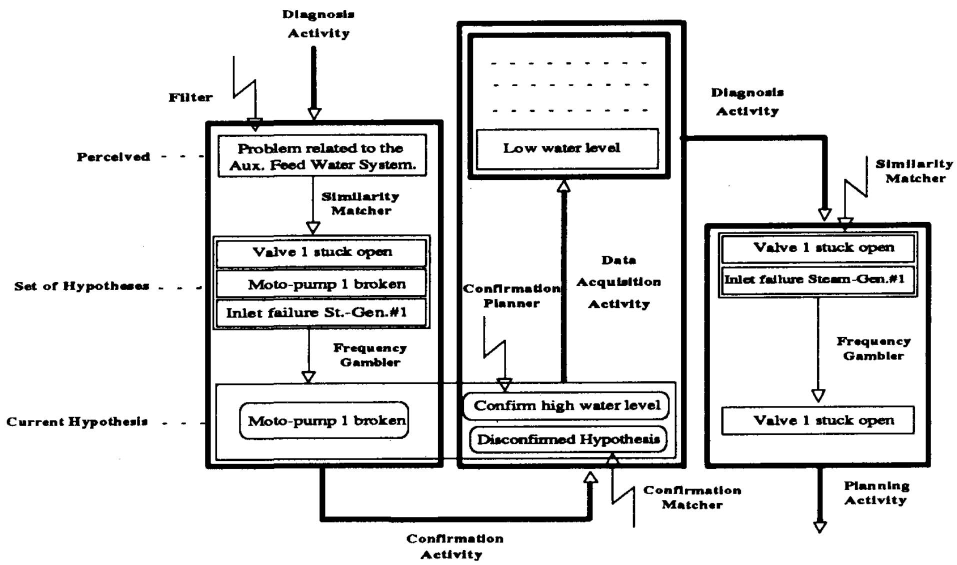


Fig. 8. Example of the diagnosis and confirmation activities.

TABLE IV
DESCRIPTION OF DBB AGENTS

Agents	Triggering	Preconditions	Body
Get-Environment	None	None	Acquisition of new set of data
Decoder	New raw data	None	Decoding new raw-data
Filter	New object on signal level	None	Filtering on basis of physical and cognitive salience
Interpreter	New object at level Sign	None	Semantical interpretation of filtered data
Similarity-matcher	New object at level perceived or new object at level CIF	None	Matching between perceived cues and knowledge in dynamic frames
Frequency-gambling	New object at level Set of Hypotheses	None	Conflict resolution on set of hypotheses
Executer	New object at level CIF	Evaluation signal activated	Execution of actions

list of numbers) to the different physical and logical variables of the plant, creating new objects on the Sign Level. Then, the Filter Agent, is applied on the sign objects in order to evaluate their expected significance. In this way, only the objects that are truly accounted for because of their overall salience in the given context are interpreted and translated to semantic values by the interpreted agent. The results of this strategy appears as objects on the Perceived Level.

The diagnosis scenario: this scenario implements the SM and FG primitives and thus the Diagnosing Cognitive Function. According to the new perceived objects, a diagnosis process can be initiated. This process takes place in two stages (Fig. 8): 1) the Similarity-Matcher Agent, triggered by new cues of the Perceived Level, generates a set of Potential Diagnoses that constitute an object of the set of the Hypotheses Level; 2) from that object, one hypothesis is selected by the Frequency-Gambler Agent. This hypothesis becomes an object of the current Hypothesis Level.

The confirmation scenario: A confirmation stage may be started by searching into the environment for pertinent information (see Fig. 8).

The confirmation strategy can be exploited to assess either a diagnosis or the result of an executing strategy. It consists of three steps: the selection of data for confirmation; the data acquisition strategy, executed on the basis of the requirement of the first step of the confirmation, which is performed by tuning the parameters of the filter agent; the actual stage of confirmation, performed by comparing expected and perceived data. Note that this scenario comprises the execution of another scenario.

Execution: When an hypothesis is confirmed, the action plan associated with the hypothesis is executed. The first action is taken from the RBF's of the operator KB, corresponding both to the current time and to the diagnosed situation .

2) A Sample Case of Man–Machine Interaction To illustrate what has been described in the previous chapter, a sample example of man–machine interaction is presented. The model simulates the various cognitive processes of an operator watching a control panel. A first set of data, indicating a nominal plant situation, is perceived by the operator who keeps on watching the panel. In the mean time the physical system evolves toward an abnormal situation. The operator detects this new situation and tries to find out an explanation. His first hypothesis is unconfirmed because of contradictory information provided by the control panel. A second hypothesis is then generated and, this time, successfully confirmed. The operator starts to execute the corresponding

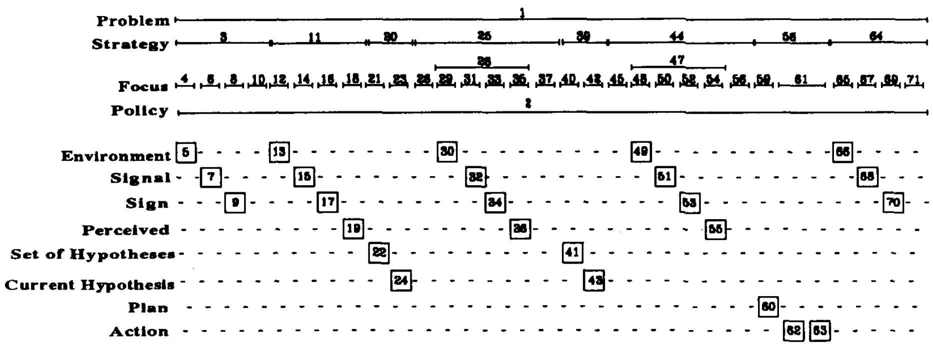


Fig. 9. History of BB after a simulation case.

recovering plan, in order to restore a stable state. The simulation ends with the perception of information indicating that the system has reached a new nominal state.

In order to simulate this behavior, the model executes 72 cycles and generates about 100 objects. Fig. 9 gives a summary of the CBB and DBB at the end of the run. Each segment represents the time interval during which a new CBB object is operative. The number on the segment is the BB-cycle of creation of the object. The DBB objects are represented by a little box containing the cycle in which they have been created. As a cycle does not necessarily produce new objects, some of them are not present on the figure.

A brief explanation of the various cycles is given hereafter.

Cycles 1–2	Put-problem-on-BB action: A new problem is put on the BB.
Cycles 3	Start-strategy action: generates the data-acquisition strategy.
Cycles 4	Initialize-focus action: generates the first focus of the current strategy.
Cycles 5	Get-environment action: Acquisition of new raw data.
Cycles 6	Update-focus action: generates the next focus of the current strategy.
Cycles 7–10	The data acquisition strategy is executed.
Cycle 11	Update-strategy. A new strategy is selected.
Cycles 12–19	Detection of an abnormal situation (Cycle 19).
Cycle 20	Update-strategy. Selection of a new strategy. A diagnosis is started.
Cycles 21–24	Execution of the diagnosis strategy.
Cycle 25	Update-strategy selects the confirmation strategy.
Cycle 26–38	Execution of the confirmation strategy.
Cycle 39–43	A new diagnosis strategy is executed.
Cycle 44–57	A new confirmation strategy is executed.
Cycles 58–63	The Executer and Planning agents are executed.
Cycles 64–71	The last data acquisition strategy is performed.
Cycle 72	Stop-problem ends the BB.

IV. Simulation of Problem-Solving Situations

A number of case studies have been developed by applying COSIMO to the simulation of a nuclear power plant operator controlling the transient behavior of a feedwater system, active only during emergencies conditions: the system of auxiliary feedwater to the steam generator (AFWS). Such system is normally in a stand-by condition and becomes automatically operative when the power station is shut down for an operational transient or in an emergency situation. The objectives of the AFWS are to supply sufficient cooling water to the steam generators during the shut-down period, when the normal cooling systems are not operative and there still exists a substantial amount of heat ("decay heat") still being produced in the plant that must be removed. The AFWS needs to be regulated at the beginning as well as during the evolution of the transient in order to adapt the heat removal capacity of the AFWS to the progressive decrease of decay heat being produced in the plant.

The initiating event considered in all sample cases has been a loss of electrical power, or "station black-out," i.e., an emergency situation. Such accident induces the nuclear reactor emergency shut-down and the consequent automatic starting of the auxiliary systems, including the AFWS. A model of the physical behavior of the AFWS has been developed accounting for a variety of faults that may occur on the plant while performing its functions as well as on the control and instrumentation system associated with the AFWS and the steam generators [20].

The AFWS and COSIMO models have been run in an interactive mode in order to create the control simulation of the operator model acting on the auxiliary feedwater system. An accident period of one hour has been simulated. This allows sufficient time for the decrease of the decay heat generated in the plant and thus for a number of control actions to be simulated by COSIMO on the AFWS.

The COSIMO input data concerning the Knowledge Base have been developed from the elicitation of an expert engineer on the management of the AFWS and by running the simulation of the AFWS for many different initial conditions and transient configuration in order to build a sufficiently rich number of RBF's for inclusion in the KB.

The remaining input data of COSIMO are the parameters, by which it is possible to tune the behavior of the main parts of the model, like, for example, the interpretation and cognitive filtering activities. The modification of these parameters, either as input data or during the dynamic simulation, allows to reproduce subtle variations in the overall model performance. A detailed description of the COSIMO parameters can be found in the Appendix. The case studies performed with COSIMO have been obtained by modifying one part of input data concerning the model parameters while maintaining untouched the other part related to the content of the KB.

In particular, four cases have been studied. The first one describes the expected behavior of the operator in the management of the AFWS, with the physical system and the control instrumentation performing optimally without failures or malfunctions during the accident evolution. The other three cases simulate further failures in the AFWS mechanical components occurring during the evolution of the accident after the initiating event of station black-out. These three cases aim at the study of deviations from the appropriate control strategy and their consequences on the overall accident management. By deviations of problem solving we intend a progressive decoupling between the plant evolving in time and the operator ability to perceive and interpret an accident, to integrate new events that modify dynamically the scenario and then to act appropriately on the system. Deviations frequently observed in dynamic and complex environments are fixations, which are failures to revise a situation assessment as new evidence is produced [40]–[42]. These fixations are difficult to identify when they occur. Sometimes, certain actions performed by an operator, or some verbal exchanges with other operators, may indicate that there is a fixation on one type of explanation for a given situation. Often, such fixations become manifest lately in the evolution of the accident, when evidence indicates that operators have not succeed in updating their views of the world.

Another type of deviations that operators are likely to show are "premature closures" [43], which result in the termination of the problem solving without generating all the alternatives and without seeking for all the available information. Moreover, a high level of workload and of stress, induced by the effects of the working environment, may also reduce the operator problem solving capacities, especially when the problems to be dealt with comprise numerous competing features, such as control actions, verification, monitoring and supervisory activities [44].

Three main aspects of deviations of problem solving have been examined in the framework of COSIMO, as sample case studies applications. They have been defined as: cognitive collapse, unadapted change and cognitive lock-up.

In summary, the four cases studied by COSIMO can be reported as follows:

A describes the *nominal* expected behavior of the model controlling the AFWS during the evolution of the station black-out accident, when no new particular events occur in the AFWS and no problem solving deviations are expected.

B describes a situation of a *cognitive collapse*, in which the operator loses the capacity of reasoning clearly. This situation is typical of sharp stress conditions, which may be generated by the complexity of the working environment or by the sudden demand of more control and supervisory activity due to new incidental events occurring during the evolution of the station black-out sequence.

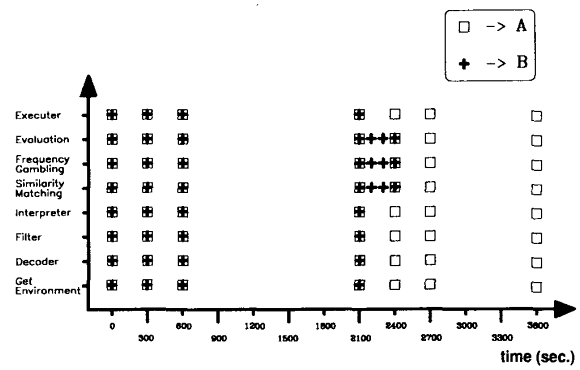


Fig. 10. Agents describing the cognitive processes in case *A* and *B*.

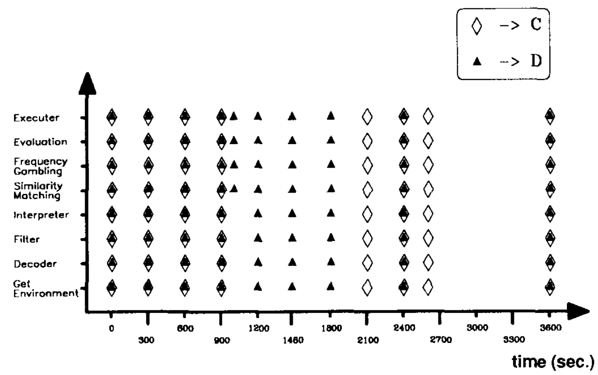


Fig. 11. Agents describing the cognitive processes in case *C* and *D*.

C describes an *unadapted change*, whereby, in presence of a new event occurring during the accident evolution, the model changes his view and diagnosis, but it fails to account for the effects of past events on the current situation.

D describes the case of a *cognitive lock-up* that causes the persistence of the first explanation even after a new plant failure has occurred during the evolution of the accident. This generates what we may call a fixation error.

In Figs. 10 and 11 the various agents describing the cognitive processes acting at different times are represented. Figures 12 and 13 indicate respectively the distributions in time of the amount of filtered variables and the similarity matching values during the four simulations. These cases and the related figures will now be described in some details.

Case A—Expected Behavior

In this case, no human error nor plant failure occur following the initiating event of station black-out. The actions and decision making of COSIMO are always appropriate and lead to a good control sequence of the decay heat removal.

The parameters assumed in the input data of COSIMO for this case are those of default value contained in the model and are shown in the Appendix. These parameter values allow the simulation of a generic operator with no effects of problem

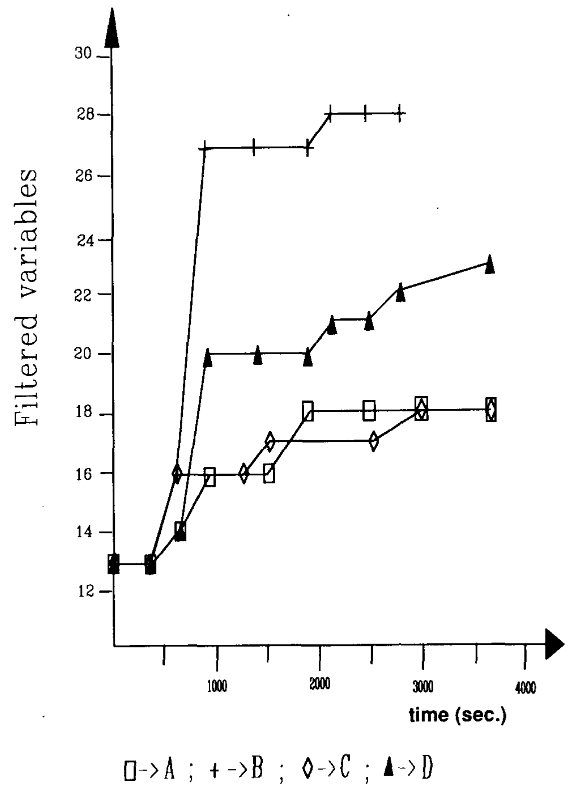


Fig. 12. Distribution of filtered variables for cases A–D.

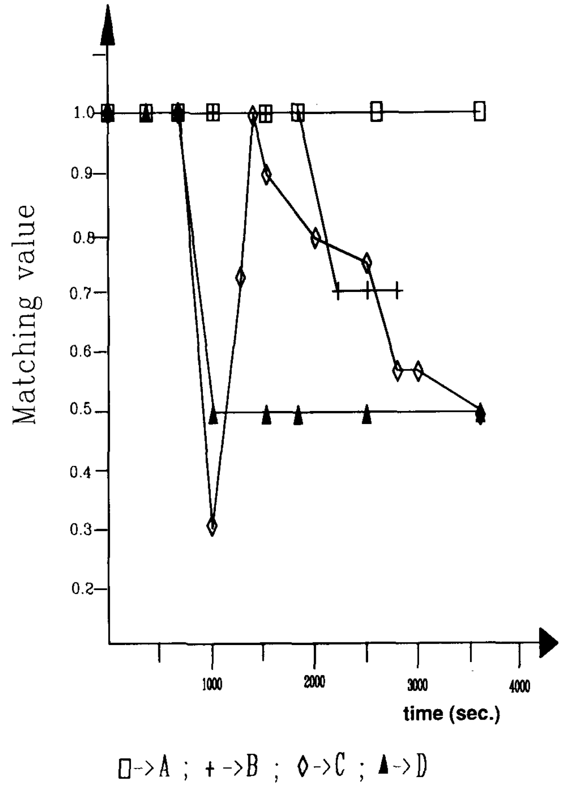


Fig. 13. Distribution of matching values for cases A–D.

solving deviations. In particular, the results of the simulation show the following characteristics:

1) the model takes into account an expected set of indicators either logical (configuration of the AFWS, position of valves, etc.) or physical (values of quantities like temperature, pressure, etc.) (Fig. 12);
2) the value of similarity matching throughout the simulation is constant and equal to 1 (Fig. 13); this means that there is a perfect correspondence, at all times, between the COSIMO expectations on the plant behavior and the cues coming from the environment, which describe and represent the effective process evolution;
3) the corresponding actions are executed according to an optimal solution of the control procedure.

Case B—Cognitive Collapse

In this case, the accident progression in the AFWS consists of the same station black-out transient as for the case A, but a further event of inlet-tube-break occurs in one of the steam generators connected to the AFWS at approximately 2100 s after the initiating event. This new event requires an immediate set of control actions in order to limit the propagation of the effects of the new failure to the other parts of the AFWS and to ensure the appropriate feeding to the remaining steam generators.

In case of cognitive collapse, the operator begins by not recognizing the new situation to be controlled and the model enters in a stage of confusion with more and more cues being filtered (Fig. 12), among others those that may not appear related to the current problem, and with no execution plan being selected and carried out (Fig. 10).

In this way, the lack of the model understanding of the current process evolution leads to a search for a larger set of indicators. Nevertheless, these extended amount of data collected (filtered) are only poorly and badly used. Indeed, when the model is in this cognitive state, it is not able to exploit further the incoming new or changed cues, which should have normally allowed a shifting in view. COSIMO does not succeed in performing an accurate diagnosis. In particular, COSIMO is well "aware" that its current view of the world is not sufficiently supported and thus that its problem representation is not a realistic one, but it does not succeed in finding the appropriate hypothesis. It tries to take into consideration all possible known explanations, jumping from one to another, without being much addressed to the current situation. It is therefore stuck in a kind of endless diagnostic stage and does not switch to the recovery activity by taking the appropriate action.

The matching value is always lower than the expectation (Fig. 13), i.e., lower than the evaluation threshold used by the evaluation agent. The number of activities called into action becomes drastically smaller in comparison to case A and they are repeated without obtaining a satisfactory resolution of the problem (Fig. 10).

In COSIMO the simulation of this problem solving deviation was obtained by:

- expanding the model field of attention, i.e., decreasing the *filter* threshold;
- decreasing the *cognitive salience* weight and increasing the *physical salience weight*;
- reducing the interpretation of small variations by increasing the *interpretation threshold*;
- increasing the *weight of the frequency* relative to the *weight of the similarity matching*;
- increasing the *evaluation threshold.*

It is interesting to observe here that, in Fig. 12, there is a noticeable difference between the set of filtered variables of case A and B even before the time 2100 s when the new event occurs. As a matter of fact, given that the two accident evolutions are the same for the first 2100 s, one may not have expected differences in the COSIMO output before that time. However, the modification of the parameters, in particular the decreased value of the filter threshold, implies that a much greater amount of cues are actually filtered by the model at all times and thus also that a much greater work of matching is actually performed by the model before reaching the maximum value 1 for the time period between 0 and 2000 s (Fig. 13).

Moreover, the considerable amount of matching analysis implies that, if the AFWS simulation had included noisy and imprecise data coming from the information system instead of precise and correct data, the tendency to the cognitive collapse may have induced a problem solving deviation even before the event at 2100 s.

Case C—Unadapted Change

The accident sequence described in this case is similar to the one of case B, but the failure event on the AFWS, i.e., the feedline break on one of the steam generator, occurs at a very early stage of the accident (300 s) and consists of a very small break of the line with delayed effects on the physical behavior of the system.

The parameters setting that has been selected in order to simulate the unadapted change of the operator during the sequence consisted in the reduction of the evaluation threshold from the default value and in the increase of the weight of the frequency versus similarity for the process of frame retrieval from the knowledge base. The remaining parameters have been kept at their default value.

As consequence of this input data configuration, at the beginning of the accident COSIMO succeeds in selecting the correct scenario through the SM and FG primitives, i.e., the plant has experienced a station black-out and the AFWS has automatically started. The second major failure event takes some time to generate relevant effects on the plant that are actually reported by the information instrumentation and noticed by the operator only with a certain delay.

This explains why the matching value remains high, i.e., 1, up to 600 s (Fig. 13), even though the failure has occurred at time 300 s. At a later time, i.e., 900 s, the new accident pattern can be objectively distinguished from the initial sequence and, at that time, the hypothesis of nominal station black-out, which represents the currently instantiated frame at 900 s, badly matches the data filtered by the model (Fig. 13). Given the mismatch, COSIMO tries to update its current view of the world by searching for a better candidate in the knowledge base, but, following the high bias of the FG primitive, COSIMO selects the wrong scenario. Consequently, the situation is missdiagnosed and the inappropriate sequence of actions is initiated. The model does not succeed in performing the expected task of appropriately cooling the system for one hour of transient.

A number of further observations can be made. Firstly, there is a substantial amount of symptomatic or surface similarity between the incorrect hypotheses selected by the model (for time 900 < seconds < 3600) and the effective state of the plant. Indeed, as an example, the similarity matching value is 1 at 1400, and it also remains high, above 0.7, between 1400 and 2100 s. This stems from the slowly evolving effects of the failure event on the AFWS and from the fact that the initial diagnosis was good and consequently the plant was properly regulated during the initial part of the accident.

The two lowest values of the matching are found at 900 s and during the final phase of the analysis. This is because up to 900 s the CIF was the initial one and no different configuration was attempted given the high value of the matching. At 900 the model "realizes" (matching value very low) that a new event has occurred and that a different configuration has to be evaluated (as it is done) and this causes an improvement in the matching value. Indeed, even if the good solution is never obtained, an initial amelioration of the matching is found.

At the end of the analysis the persistent failure in diagnosing correctly the situation leads to a progressive degradation of the plant and therefore to a continuous decrease of the matching value. The misdiagnoses at all times result from the slight difference in frequency between a number of possible cases, which favors the selected erroneous frame. The selection of the most effective schema of the basis of a frequency criterion, which compensates for the difference in matching, is a commonly encountered and proven deviation of cognition in accidental circumstances such as the one considered here [15]. Moreover, the low value of evaluation threshold allows a weakly supported hypothesis to be trusted even in presence of a decreasing matching with the cues of the environment.

Case D—Cognitive Lock-Up

In a cognitive lock-up condition, the model is fixed on a selected hypothesis and continues acting on the system in the light of an obsolete explanation even in presence of strong disconfirming information. A progressive de-coupling is then observed between the model understanding and the evolution of the plant.

The accident scenario simulated in this case is totally similar to the previous case and, thus, as explained above, the failure event on the AFWS takes some time to manifest its consequences on the information system because of the slow dynamics of the process. Consequently, as in the previous case, the similarity matching of the CIF "nominal station black-out" with the actual cues of the environment remains at level 1 up to approximately 600 s. At this time is not yet possible to acknowledge that a new failure has occurred.

At 900 s the accident patter of the CIF does not satisfy the actual plant evolution, but COSIMO does not succeed in noticing the change from the indications of the AFWS. In this case, the nominal station black-out accident scenario, on which COSIMO is fixed, has a greater subjective frequency than the inlet break scenario, which is the dominating event of this sequence. Even if the abnormal behavior of AFWS is demanding a greater cognitive effort of variables collection and interpretation (Fig. 12), the high satisfaction of the model about its current hypothesis does not allow any attempt to update the reasoning in the light of the disconfirming cues. This leads to a poor matching of the selected cues with the expectation (Fig. 13). Moreover, even with a very low similarity matching value, the model remains confident in its explanation and does not succeed at revising the situation assessment. Therefore, it carries out the initially selected strategy, step by step, according to the scheduling of actions contained in the frame with no attempt to perform a revision.

The sequence of control actions is obviously not successful at the end of the one hour transient simulation.

Technically, this fixation pattern has been obtained mainly by modifying two parameters of the model: the weight of the frequency gambling and the evaluation threshold. The weight of the frequency gambling has been increased with reference to the similarity matching for the computation of hypotheses. This gives more importance to the past experience than to the observation of actual cues of the world.

The evaluation threshold has been increased. This has made the model less careful about possible mismatches between the hypothesis driven expectations and the interpreted cues of the world. Therefore, even if the matching value becomes very low from 900 s onward (Fig. 13), the support to the initial CIF "AFWS nominal station-black-out" remains substantially high and, in any case, sufficient for the initial hypothesis to be trusted all throughout the sequence.

V. Conclusion

Cognitive modeling is surely a complex and ambitions enterprise, specially when dealing with process plant transient that have an intrinsic complexity of their own concerning physical processes, control procedures and variety of scenarios to be dealt with. However, the current state of information technology tools, the necessity to consider cognitive aspects in studies of safety of plants and the high level of mutual understanding reached between engineers, ergonomists and psychologists are clear symptoms that such type of modeling can be developed and can be usefully applied, at least, within well-defined working situations. The model COSIMO presented in this paper is certainly the results of such type of need and synergism.

The first phase of development has resulted in a model that is currently addressing, with particular attention, the problem of human decision making and behavior during accident management. In particular, in this paper two aspects of the modeling architecture have been discussed in detail, namely:

- the formulation of a theoretical framework, which has only partially been exploited and expanded beyond the identification of the fundamental mechanisms of cognition; and
- the instantiation of the modeling paradigm in a blackboard computational system, which offers a very rich possibility of flexibility and expansion, needed once the theory will be further developed.

The identified limitations, relative to the absence in the simulation of the processes of planning driven by the direct inference primitive, are the current major theoretical concern of the authors. Indeed, the analysis of the actual performances of real operators during the management of the accidents [45] has demonstrated that, while on the one hand, the two primitives of cognition of SM and FG are, in most cases, predominant in the decision making processes, on the other hand however, it has shown that a substantial inferential reasoning is performed by the operators, even for the scenarios where a strong effect of SM and FG was reported. In particular, it has been shown that when carrying out a selected pattern of actions, a continuous process of mental work was performed concerning the state and the dynamic evolution of the plant under control.

This latter point implies that, as far as the future development of COSIMO is concerned, the theory must be expanded, to include inferential reasoning with special models able to handle peculiar cognitive processes, such as temporal and analogical reasoning, use of metaknowledge, strategical thinking etc.

Nevertheless, the model has been shown to be able to simulate a number of deviations from nominal behavior that have resulted in erroneous actions and decisions and in inappropriate management of the accident. Concerning this subject, the architecture of COSIMO needs to be further improved with a consistent theoretical formulation of the causes and the manifestations of erroneous behavior [46] in order to refine the definition of the input parameters and the interface/interaction of the cognitive model with the plant behavior model.

The results shown in the paper on the analysis of the consequences of these deviations have also proven the ability of the model to cover a number of the expected applications of COSIMO in a context of simulation of human–machine interaction, as it was discussed in the introductory section of this paper. In particular, given a specific plant configuration and control design, COSIMO:

- has evaluated the accident evolution and the consequences of both expected normative and inappropriate behaviors;
- has analyzed how an operator is likely to act, given a particular event, namely the occurrence of a second failure following the major accident initiator (sample cases B, C, and D); and
- has identified difficult problem solving situations, such as the cases of plant failures with slowly appearing effects, in presence of limited or bounded cognitive resources (sample cases C and D).

The consideration of the entire spectrum of potential applications of COSIMO, in particular for the identification

TABLE V
COSIMO PARAMETERS AND DEFAULT VALUES

Parameters	Definitions	Default Values
Filter threshold	Threshold value for the salience of cues in order to gain acceptance through the filter	0.75
Physical salience weight versus Cognitive salience weight	Relative influence of physical versus cognitive factors in the computation of the salience of a cue	2
New cue weight Changed cue weight Symptom weight	Weights used to qualify either new cues firstly appeared, or changed cues or persistent cues	1 0.9 1
Cognitive inhibition	Set of quantities to be inhibited	nil
Interpretation threshold	Sensitivity (relative versus absolute) in accounting for the variation of quantities to be interpreted	10/60
Similarity weight Frequency weight	Weight of SM and FG primitives in selection of hypotheses	1 1
Extended search Time search	Amount of KB and time frames considered in selecting and asserting a hypothesis	All KB 0
Evaluation threshold	Degree of confidence for a hypothesis	0.7

of situations leading to erroneous behavior, for the analysis of recovery and for the study of the effects of changes in the man–machine systems, requires the availability of models of inferential reasoning, by which the simulation of unfamiliar and unknown situations can be handled. The use of analogical reasoning and inferential mechanisms are examples of theoretical approaches that will be attentively considered in the future development of COSIMO.

ACKNOWLEDGMENT

The authors would like to thank James Reason for his essential collaboration in the conceptual development of the cognitive model and in many sessions of discussion. Thanks is also due to G. Volta and G. Mancini for their continuous encouragement and precious advice in the performance of this work.

APPENDIX
DESCRIPTION OF PARAMETERS

Table V shows the parameters that are used for defining and selecting the sequence of the cognitive activities. These parameters are the input data of COSIMO, by which intraindividual differences and special deviation mechanisms of problem solving are simulated. A set of defaults values for these data is built-in the program by which the behavior of an average subject is matched. The other data refer to the construction of the KB, which must be done for each case of human–machine interaction and by which the expertise of different subjects is defined.

Filter threshold: The filter threshold governs the attentional level of the model. Decreasing the filter threshold extends the field of attention and thus more cues of the environment are collected. This does not mean that only pertinent cues are taken into account: indeed, the subject can be, for instance, focused on unimportant details.

Increasing the filter threshold means to narrow the field of attention by which only very salient cues are taken into account. The selection of information for interpretation depends also on the salience content of each cue.

Cognitive versus physical salience: The salience is the parameter that measures the attention grabbing power of a cue. Two types of salience are defined in the model: the cognitive salience and the physical salience, which refer respectively to the subjective significance and the physical properties of cues.

Greater importance can be given to the physical or to the cognitive salience, by adjusting the relative weights assigned by the two parameters "physical salience weight" and "cognitive salience weight".

New and changed cues: New and changed cues describe respectively the new events occurring during a dynamic human–machine interaction and the cues that modify their values during the evolution in time. In COSIMO, different weights can be given to newly appeared cues or to changes that have occurred. At certain moment it can be crucial for the operator to interpret new information that are arriving, while in other circumstances it may be more important to be able to interpret changes in the system.

Cognitive inhibition: The cognitive inhibition is a parameter by which part of the system may be inhibited from reaching attentional level. For example when a part of a system does not work because of a previous incident or because it has been put out of the loop for a maintenance, the operator knows that this system can be neglected. It can be said that this part of the system is, from the operator point of view, cognitively inhibited and other events related to it are not taken into account.

Cognitive inhibition can be a very economic process because it allows to neglect some unimportant or not pertinent events and to limit the information search at certain moment. But it can be source of errors too.

Interpretation threshold: The interpretation threshold regulates the sensitivity of the model to changing cues. The greater the interpretation threshold, the greater is a variation of a reading on the control panel that is interpreted as a stable state.

The interpretation threshold is assigned in percentage or fraction, with reference to the value of the observed cue at the previous time step.

Similarity versus frequency: Similarity matching and frequency gambling are the two main primitives of cognition that govern the decision making of the model.

Differential weights can be given to the similarity matching or to the frequency gambling in order to privilege either of the two primitives. During the simulation of the human–machine interaction, each frame selected from the knowledge base shows a matching value and a frequency tag: the final selection is based on a weighted combination of these two factors. Consequently, the relative weights are assigned as input for simulating different tendencies of human behavior.

Extended and temporal search: These two parameters are used to improve the dynamic search process in the knowledge base. The first parameter, i.e., extended search, delimits the amount of the knowledge base, in terms of the number of frames, over which the search is performed for the selection of the CIF.

The second parameter, i.e., temporal search, defines the amount of time history or time periods that can be revised for information retrieval and/or for evaluating trend indications.

Evaluation threshold: The evaluation threshold specifies the lowest support value that can be associated to an hypothesis in order to be trusted and accepted by the model. The higher is the value of the evaluation threshold, the more cautious or even suspicious becomes the simulated behavior.

An evaluation stage aims at deciding if a potential explanation can be trusted or if it must be rejected. The hypothesis selected has the highest support, but this can be too low to give the model any confidence in this explanation.

REFERENCES

[1] L. Bainbridge, "The ironies of automation," *Automatica*, vol. 19, no. 6, pp. 775–780, 1983.

[2] G. Mancini, "Modeling humans and machines," in *Intelligent Decision Support in Process Environments*, NATO ASI Series, E. Hollnagel, G. Mancini and D. D. Woods, Eds. Berlin: Springer-Verlag, 1986.

[3] T. B. Sheridan, "Forty-five years of man-machine systems: History and trends," keynote address in *Proc. 2nd IFAC Conf. Anal., Design and Evaluation of Man-Machine Syst.*, Varese, Sept. 10–12, 1985.

[4] N. Moray, "Monitoring behavior and supervisory control," in *Handbook of Perception and Human Performance*, K. Boff, L. Kaufman, and J. Thomas, Eds. New York: Wiley, 1986.

[5] E. Hollnagel, "What do we know about man-machine systems," *Int. J. Man Machine Studies*, vol. 18, pp. 135–143, 1983.

[6] G. Apostolakis, P. Kafka, and G. Mancini, "Accident sequence modeling: Human actions, system response, intelligent decision support," special issue of *Reliability Engineering and System Safety*, vol. 22, 1988.

[7] V. De Keyser, "Cognitive development of process industry operators," in *New Technology and Human Errors*, J. Rasmussen, K. Duncan and J. Leplat, Eds. New York: Wiley, 1987.

[8] D. D. Woods, E. M. Roth, and H. Pople, "Cognitive environment simulation: An artificial intelligence system for human performance assessment," *Modelling Human Intention Formation*, vol. 2, Nureg-CR-4862, Washington DC, 1987.

[9] K. S. Rubin, P. M. Jones, and C. M. Mitchell, "OFMspert: Inference of operator intentions in supervisory control using a blackboard architecture," IEEE Transactions on Systems, Man and Cybernetics, Vol. 18, 4, 618–637,1988.

[10] R. Amalberti, M. Bataille , F. Deblon, A. Guengant, J. M. Paignay, C. Valot, and J. P. Menu, "Développement d'aides intelligentes au pilotage: Formalisation psychologique et informatique d'un modèle de comportement du pilote de combat engagé en mission de pénétration," Rep. "CERMA" 89.09, 1989.

[11] L. Pinsky and J. Theureau, "Signification et action dans la conduite de systèmes automatisés de production séquentielle," *Collection d'Ergonomie et de Neurophysiologie du Travail*, Paris: CNAM, no. 83, 1985.

[12] H. G. Stassen, G. Johannsen, and N. Moray, "Internal representation, internal model, human performance model and mental workload," *Automatica*, vol. 26, no. 4, pp. 811–820, 1990.

[13] W. Rouse, "Models of human problem solving: Detection, diagnosis and compensation for system failures," *Automatica*, vol. 19, no. 6, pp. 613–625, 1983.

[14] J. Rasmussen, *Information Processes and Human-Machine Interaction—An Approach to Cognitive Engineering*. Amsterdam: North Holland, 1986.

[15] J. Reason, *Human Error*. Cambridge, MA: Cambridge Univ. Press, 1990.

[16] P. C. Cacciabue, G. Mancini and U. Bersini, "A model of operator behavior for man-machine system simulation," *Automatica*, vol. 26, no. 6, pp. 1025–1034, 1990.

[17] B. Hayes-Roth, "A blackboard architecture for control," *Artificial Intell.*, vol. 26, no. 2, pp. 251–321, 1985.

[18] E. Engelmore and T. Morgan, *Blackboard Systems*. New York: Addison-Wesley, 1988.

[19] J. P. Nordvik, P. C. Cacciabue, F. Decortis, and M. Masson, "Implementation of a cognitive model in a blackboard architecture," 8th Eur. Annu. Conf. Human Decision Making and Manual Contr., Copenhagen, 12–14 June 1989.

[20] P. C. Cacciabue, A. Codazzi, and F. Decortis, "The analysis of the functional role of man and machine in the control of a notional auxiliary feedwater system," CEC Rep., EUR 13172 EN 1990.

[21] J. Reason, "Recurrent errors in process environments: some implications for the design of intelligent decision support systems," in *Intelligent Decision Support in Process Environment*, NATO ASI Series, E. Hollnagel, G. Mancini, and D. D. Woods, Eds. Berlin: Springer-Verlag, 1986.

[22] J. Reason, "Cognitive aids in process environments: Prostheses or tools?" *Int. J. Man-Machine Studies*, vol. 27, nos. 5 and 6, pp. 463–471, 1987.

[23] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communications*. Urbana, IL: Univ. Ill. Press, 1949

[24] K. Krippendorf, "Information theory—Structural models for qualitative data," in *Quantitative Applications in the Social Sciences*. A Sage Univ. Paper, 62, Beverly Hills, CA: Sage, 1986.

[25] S. Vosniadou, A. Ortony, Eds., *Similarity and Analogical Reasoning*. Cambridge, MA: Cambridge Univ. Press, 1989.

[26] A. M. Aitkenhead and J. M. Slack, Eds., *Issues in Cognitive Modeling*, LEA, Open University set book, London, 1990.

[27] J. Rasmussen, "Outlines of a hybrid model of the process operator," in *Monitoring Behavior and Supervisory Control*, and T. Sheridan and G. Johannsen, Eds. New York: Plenum, 1976.

[28] G. Iosif, "La stratégie dans la surveillance des tableaux de commande—II: Quelques aspects de l'activité de surveillance chez les opérateurs dans la production semi-automatisée," *Rev. Roum. Sci. Soc-Psychol.*, vol. 12, no. 1, pp. 75–94, Bucarest, Romania, 1969.

[29] E. Edwards and F. Lees, *The Human Operator in Process Control*. London: Taylor & Francis, 1974.

[30] V. De Keyser, F. Decortis, and A. Housiaux, and A. Van Daele, "Les communications hommes-machines dans les systèmes complexes," *Action Fast*. Bruxelles: Politique Scientifique, 1987.

[31] L. Bainbridge, "Cognitive task analysis for process operations. A model of cognitive processes and its implications," Workshop Cognitive Processes in Complex Tasks, Wilgersdorf, RDA, 4–6 December 1989.

[32] R. J. Beishon, "An analysis and simulation of an operator's behavior in controlling continuous baking ovens," in Edwards and E. F. Lees, Eds., *The Human Operator in Process Control*. London: Taylor & Francis, 1974.

[33] J. Rasmussen and A. Jensen, "Mental procedures in real life tasks: A case study of electronic trouble shooting," *Ergonomics*, vol. 17, pp. 293–307, 1974.

[34] J. Reason, "Modeling the basic error tendencies of human operators," *Reliability Engineering and System Safety*,vol. 22, pp. 137–153, 1988.

[35] L. Bainbridge, "Forgotten alternatives in skill and workload," *Ergonomics*, vol. 21, no. 3, pp. 169–185, 1978

[36] J. Leplat, *Erreur Humaine, fFiabilité Humaine dans le Travail*. Paris: Armand Collin, 1985.

[37] E. Hollnagel, "Modeling of cognition: Procedural prototypes and contextual control," in preparation for *Int. J. Man-Machine Studies*.

[38] R. Amalberti and F. Deblon, "Cognitive modeling of aircraft process control: A step torwards an intelligent onboard assistance," *Int. J. Man-Machine Studies*, vol. 36, no. 6, pp. 639–671, 1992.

[39] H. P. Nii, "Blackboard systems (Part 1)," *AI Mag.*, vol. 7, no. 2, pp. 38–53, 1986.

[40] V. De Keyser and D. D. Woods, "Fixation errors in dynamic and complex systems," in A. G. Colombo and R. Misenta, Eds., *Advanced Systems Reliability Modelling.* Dordrechts, The Netherlands: Kluwer, 1989.

[41] M. Masson, "Understanding, reporting and preventing human fixation errors," in T. W. Van der Schaaft, D. A. Lucas, and A. R. Hale, Eds., *Near Miss Reporting as a Safety Tool.* Oxford, UK: Butterworth-Heinemann, 1991

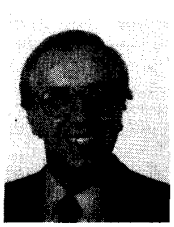
[42] D. D. Woods, "Some results on operator performance in emergency events," in *Ergonomic Problems in Process Operations.* Inst Chem. Eng. Symp. Ser. 90, D. Whitfield, Ed., 1984.

[43] I. L. Janis, "Decision making under stress," in *Handbook of Stress. Theoretical and Clinical Aspects*, L. Goldberger and S. Breznitz, Eds. New York: Free Press, 1983.

[44] N. Moray, "Mental workload its theory and measurement," *NATO Conference Series, III.* New York: Plenum, 1979

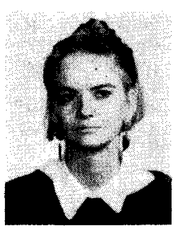
[45] A. Bellorini, P. C. Cacciabue, and F. Decortis, "Validation and development of a cognitive model by field experiments: Results and lesson learned," in *Proc. 3rd Eur. Conf. Cognitive Sci. Approaches to Process Control*, Cardiff, Wales, 2–6 Sept., 1991.

[46] E. Hollnagel, "The phenotype of erroneous actions: Implications for HCI design," in G. R. S. Weir and J. L. Alty, Eds., *Human Computer Interaction and the Complex Systems.* London: Academic, 1991.



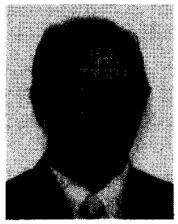
Pietro Carlo Cacciabue was born in Nizza Monferrato (Asti), Italy, on November 4, 1949. He graduated in nuclear engineering at the University of Turin, Turin, Italy, in 1973

Since 1975 he has been a researcher of the Commission of the European Communities. His first position was at the experimental "High Temperature Gas Cooled" nuclear power plant in Winfrith-UK. In 1976, he entered the Joint Research Centre (JRC) at Ispra and was engaged in theoretical studies and in reliability and probabilistic safety assessment analysis of water reactors. Since 1982, his major interest has been in the study of man–machine systems, with particular emphasis on the development of models of human behavior to be included in safety evaluations as in decision support tools. Since 1991 he has been responsible for the research activity in the field of working environment within the Institute of Systems Engineering and Informatics of JRC.



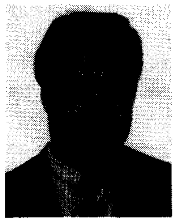
Françoise Decortis received the Ph.D. degree in psychology in 1992 from the University of Liège, Liège, Belgium.

She was a Research Assistant at the Department of Work Psychology at the University of Liège from 1984 to 1986. She joined the Institute for Systems Engineering and Informations, Commission of the European Communities, in 1987. Her research interests include the study of cognition in real situations, operator's problem solving in remote control tasks, and temporal reasoning. She performed extensive field research in working environments in nuclear power plants, blast furnaces, and air traffic control.



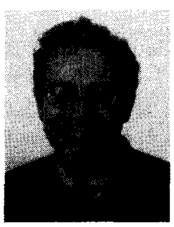
Bartolome Drozdowicz was born in Santa Fe, Argentina, in May 1950. He graduated in 1976, obtaining an electrical and electronic engineering degree from the University of Cordoba, Cordoba, Argentina.

In 1977 he joined the Department of Physics of the Cordoba University. Since 1980 he has been working at the Department of control of the Development and Research Institute INGAR as a Researcher of the National Research Council of Argentina (CONICET). During 1990 he was a Visiting Scientist at the Institute of Systems Engineering and Informatics of the Joint Research Centre, Ispra Establishment, Commission of the European Communities. His fields of interest include human reliability, cognitive simulation, real time process control, training simulation.



Michel Masson was born in Liege, Belgium, on November 4, 1962. He graduated in psychology at the University of Liège in 1985.

He joined the Department of Psychology there, first as a Researcher, and then as an Assistant Professor in 1987. In 1988–1989 he was a Ph.D. grant-holder of the Commission of the European Communities, performing a study on evaluating and modelling human errors in complex working environments. In 1990, he joined Aerospatiale Protection Systemes (APSYS, France), as a human factors specialist, where he applied his background knowledge in real domains, such as public transport and avionics. In 1991 he returned to the University of Liège, where he is currently completing his work on his Ph.D. dissertation. His major interest are in the study of human behavior modelling, human error manifestations and in computer implementation of cognitive theories.



Jean-Pierre Nordvik was born in Brussels, Belgium, in 1963. He received the masters degree in electromechanical engineering in 1985 from the University of Brussels (ULB).

From 1985 to 1989 he was a Research Assistant at the Artificial Laboratory of this University (IRIDIA). He is now Scientific Officer at the Institute of Systems Engineering and informatics, Joint Research Center, Commission of the European Communities. His research interests include distributed intelligent systems, approximate reasonings, and genetic algorithms applied to safety and reliability of man–machine systems.