1. INTRODUCTION

The influence of the so-called human factor on the safety of transportation and complex industrial systems, nuclear power plants (NPPs), and hazardous chemical installations in particular, is widely recognized. The human system or, simply, human interactions (HIs) is the term that describes all interfaces between humans and the system (Moieni, Spurgin, & Singh, 1994a, 1994b). Due to severe accidents that have occurred in industrial and transportation systems, for example, the Three Mile Island NPP (1979), the Chernobyl NPP (1986), the Bhopal chemical plant (1984), the Challenger shuttle (1986), the Piper Alpha oil platform explosion, and the capsize of the Herald of Free Enterprise (1988), and other accidents (Reason, 1990), the scientific community is convinced that the influence of human factors on the safety of technical systems is very significant, and often crucial. Thus, errors committed by man managing, operating, and maintaining these systems are often the most significant causes of accidents and risks associated with their operation (Dougherty & Fragola, 1988).

On the other hand, the state of the art of the human reliability methodology and the currently available methods and techniques for the probabilistic assessment of human failures indicate that this methodology is not mature. This has been confirmed by the results of some experimental research, aimed at the validation of the human reliability analysis (HRA) models, more frequently used in engineering practice, and creating doubts and discomfort among some researchers (Dougherty, 1990). Another problem is associated with the fact that the human reliability assessments are profoundly dependent on expert judgment. Results of HRA
benchmark exercises, summarized, for example, in Pouzet (1988), have shown that the human error probability (HEP) and assessed frequency of some accident situations for a specified plant equipped with safety-oriented systems, obtained by different groups of HRA experts, can reach discrepancies as high as orders of magnitude.

HRA in the context of the probabilistic safety analysis (PSA) is an attempt to model HIs and predict the impact of such interactions on the reliability and safety of plants. When a system is complex with a large number of human interactions in the various phases of the plant's normal or abnormal operations, then HRA becomes an extremely important part of PSA for a realistic assessment of plant safety (Moieni et al., 1994a, 1994b).

Therefore, considerable research efforts have been undertaken at some research institutions and regulatory bodies to improve the methodology, verify models, and propose approaches standardizing HRA performed within PSA studies (International Atomic Energy Agency [IAEA], 1992). This is very important for obtaining correct quantitative results, often used in engineering practice for reliability management and safety-related decision making. For accomplishing these objectives, a tendency of growing interest in computer-aided PSA and HRA is observed.

This article addresses some methodological and practical issues of HRA in the context of probabilistic safety studies. An integrated approach for performing PSA and HRA using the expert system technology is outlined. At the end, some current research challenges in the domain are discussed.

2. HUMAN RELIABILITY ANALYSIS FRAMEWORKS AND MODELS

2.1. Human Performance and Reliability

Hardware reliability analysis is supported by a mathematical theory, the techniques of data collection and its analysis, and specific engineering methods. In principle, the hardware reliability deals with failures over a mission (while operating), but in cases of standby or protecting systems, other types of failures have been also observed. Thus, hardware failures modeled in engineering practice can be roughly classified as (a) failures while operating, (b) failures while in standby, and (c) failures on demand. The first two failure types occur over time but the last type is rather time independent, occurring at a specific point in time. Some analogies between hardware and human reliability are sometimes perceived (Dougherty & Fragola, 1988). The range and source of environmental influences (stresses and shocks) are typically much greater for people than for equipment. People can fail during “normal” activity. Also, a situation can “demand” human action to intervene or mitigate its consequences. In contrast to hardware, normal human activity rather leads to time-independent human failures (slips), whereas the event-demanded activity is more susceptible to time-dependent failures (mistakes).

The mathematics used to evaluate the human reliability is different from that of hardware. Unlike the hardware time-dependent model, the random variable is the elapsed time for a successful response, \( T \). In this case, the reliability function can be expressed as:

\[
R(t) = \text{Pr}[T \leq t]
\]

that is, the probability that a success is achieved by time \( t \). CDF (cumulative distribution function) \( F(t) = R(t) \) is the probability that a successful response will occur at least by time \( t \). \( 1 - F(t) \) is the nonsuccess or failure probability at time \( t \). This complementary cumulative distribution function (CCDF) is used in some human reliability models as a time reliability correlation (Dougherty & Fragola, 1988).

In the evaluation of human reliability, there are also justifications to assume, for some cases, the human reliability (probabilistic) models with time-independent failures. Generally, a time-independent failure occurs following the establishing of a goal and making of a plan to achieve this goal, that is, the goal is the demand. At this point, habit or procedure usually takes over to implement the plan with possibilities to commit slips or lapses (see Section 3). Such slips and lapses manifest themselves in one of three main ways (Dougherty & Fragola, 1988): (a) omission—a step in the plan is not performed, (b) mis-selection—in performing a step, the
wrong object is used or manipulated, and (c) repetition—a step is performed again unnecessarily. The probability (a measure of the human reliability) for time-independent failures is estimated as the number of failures of a given type divided by the number of demands that could have led to that type of failure.

Human reliability is understood in this article as a quality of human performance interacting within a complex system. A measure of human reliability is the probability that a person will perform adequately specified tasks or more complex supervising and control functions for the situation considered. Two distinct cases of human failures are distinguished: time-dependent and time-independent. The purpose of the human reliability analysis (HRA) is a qualitative evaluation of the error-prone situation and the quantitative probabilistic assessment of human failure. The probabilistic assessment can be based on the results of experiments (e.g., on simulators) or using the appropriate method, selected from HRA quantifying techniques, developed during the last two decades.

2.2. Human Reliability Data and Quantifying Techniques

Human reliability data has been collected mainly in the United States since the 1960s (Dougherty & Fragola, 1988). Earliest interest in the man–machine interface was mostly confined to the military area and the first data sets reflected this perspective. As the nuclear power industry sought to introduce quantitative estimates of human performance into their safety analyses in the 1970s, expert judgment was used to extrapolate the existing body of military experience to commercial nuclear plant operation. It resulted in the publication of a handbook of THERP (Swain & Guttmann, 1983), which codifies this extrapolation. Lately there have been some efforts to create human reliability databases through an evaluation of raw performance data from the industry and experiments carried out on laboratory stands or simulators (Moieni et al., 1994a, 1994b). Such trials are known and some compiled data exist. The way from raw data to useful information is, however, laborious and biased with human judgment (Dougherty & Fragola, 1988; Gertman, Blackman, Haney, Seidler, & Hahn, 1992), especially for cases of more cognitive human actions and following errors.

Parallel research efforts have been undertaken to develop human reliability models, which are more general and applicable for some defined attributes of the situations analyzed. The first human reliability methodologies were dominated by the mechanistic view of human performance. In this behavioristic view, only the consequences of human errors are accounted for, without considering the reasons and the underlying mechanisms of these errors (Dougherty & Fragola, 1988). From the middle of the 1970s, some human reliability modeling frameworks were proposed on the basis of the achievements of the psychological theory of cognition and error psychology (Rasmussen, 1983; Reason, 1990). Unfortunately, some of the models proposed have not been verified on the basis of field or experimental data.

Thus, some limited data sources and various HRA techniques are now available. The problem is how to apply them in the analysis process, for a specific man–machine interface, to cover a wide spectrum of possible human interactions at different cognitive levels with underlying error mechanisms, and to identify more important human induced contributors to the risk.

Table 1 presents three groups of the HRA methods and techniques that are frequently used in the PSA practice. Access to source publications for some of these techniques is difficult, but there are several papers and reports containing more or less detailed descriptions with assessments of quality and usefulness in engineering practice (e.g., Cacciabue, 1988; Humphreys, 1988; Kosmowski, Degen, Mertens, & Reer, 1994).

2.3. Incorporation of Human-Induced Failure Events Into PSA Logic Structures

An interesting framework for incorporating human–hardware interactions into PSA, called Systematic Human Action Reliability Procedure (SHARP) was developed by Hannaman and Spurgin (1984). According to this procedure, the responsibility for incorporating human interactions into the PSA logic structure (usually the fault and event trees) is shared between the
system analysts and the human reliability analysts. The objective of this framework is to help the analysts in determining interactions that are important to the risk.

The SHARP framework matured in the form of seven steps. Each step has its defined objectives, inputs and outputs, activities, and rules. The links between these steps and key decision points of the procedure are shown in Figure 1. The general goals for each step are as follows:

1. **Definition**—To ensure that all the different types of human interactions are adequately considered in the study.

2. **Screening**—To identify the human interactions that are significant to the operation and the safety of the plant.

3. **Breakdown**—To develop a detailed description of important interactions by defining the key influence factors necessary to complete the modeling. The human interactions modeling consists of a representation, impact assessment, and quantification.

4. **Representation**—To select and apply techniques for modeling important human interactions in logic structures. The techniques selected should help to identify significant human actions that might impact the system logic trees.

5. **Impact assessment**—To explore the impact of significant human actions identified in the preceding step on the system logic trees.

6. **Quantification**—To apply appropriate data or quantification methods to assign probabilities for the various interactions examined, to make the sensitivity analysis and assess uncertainty ranges.

7. **Documentation**—To include all necessary information for the assessment to be traceable, understandable, and reproducible.
For the situations of more cognitive human performance, the construction of the operator action tree (OAT) is useful (Hannaman & Spurgin, 1984) to illustrate, in its consecutive sequences, the situation context and the dependencies of human actions and failures (see Section 4.2).

A new framework called SHARP 1 was developed (Wakefield et al., 1992) to enhance the process of performing HRA in a PSA via inclusion of experience and insights gained over the last several years since the SHARP framework was introduced. The SHARP 1 framework is organized into four stages (Moieni et al., 1994a, 1994b) to more closely parallel the major PSA steps:

1. HIs event definition and integration into the plant logic model. The purpose of this stage is to identify and define the various types of HIs and to describe how these interactions are incorporated into the PSA logic models.
2. HIs event quantification. This stage is aimed at quantifying the HIs that are defined in Stage 1 and it provides as output HEP for each of them.
3. Recovery analysis. At this stage the analyst uses the interim accident sequence frequency results as the input information. Recovery actions are identified for important scenarios and judged to be feasible and explicitly defined. Revisions to the PSA plant logic models are then performed for the integration of newly defined recovery actions. The final PSA accident sequence quantification is then performed and the results documented.
4. Internal review. This stage addresses the need for internal project review of the products from the first three stages and provides a final check on the development, internal consistency, and documentation of results.

SHARP 1, in particular, provides guidance to the HRA analyst on how to define the human interaction events, their boundaries, and appropriate inclusion into the plant logic models. Special attention is given to the identification and treatment of various types of HI dependencies.

2.4. Approaches Based on Integrated Simulation or AI Methods

Several new techniques have been developed for modeling the cognitive behavior of the operators during abnormal situations that occur in dynamic process systems. These include, for example, the Cognitive Environment Simulation (CES) concept (Woods, Pople, & Roth, 1990), an integrated approach for modeling the human and the machine, which has been proposed at JRC-Ispira based on the DYLAM technique (Cacciabue, 1988), and a model named COSIMO (COgnitive SImulation MOdel) for modeling the human operator controlling a complex environment (Cacciabue, Decortis, Mancini, Masson, & Nordvik, 1989).
These methods are not directly useful for the quantification of HEPs. They were developed rather for different purposes, for example, to evaluate the quality of man–machine interface and the influence on human performance in dynamic situations of the plant of such factors as stress, the quality of operator training, the quality of procedures, and so forth. Some of these new methods are related to the psychological theories of cognition and AI methods, and are aimed at modeling the human operator as a cognitive filter, a diagnoser, a planner, and so forth. There are opinions that cognitive modeling is a fundamental issue for future more advanced human reliability analysis in event-driven situations (Cacciabue, 1992).

2.5. Practical and Quality Aspects of HRA
Several of the HRA techniques available at present are more frequently used in engineering practice (Humphreys, 1988). The selection and application of HRA techniques for the situation considered is not always obvious and, therefore, is susceptible to the subjective interpretations of analysts. This has been confirmed by the results of the already mentioned HRA benchmark exercises summarized in Poucet (1988).

Thus, the problem arose of how to help the analysts select appropriate technique(s) and carry out the analysis for the situation of interest. The general idea can be to propose a guideline with computer programming implementation as an expert system. Such a software system should help the analyst to select the appropriate method on the basis of some attributes of the situation analyzed, to support the conducting of analyses and calculations, and—finally—to document important assumptions, expert opinions, and the results obtained. A concept of such a system is outlined in Section 4.

3. HUMAN PERFORMANCE, ERRORS, AND SOME FACTORS CONSIDERED IN HRA

3.1. Human Behavior Types
The distinction of three categories of human behavior was proposed by Rasmussen (1983). His conceptual framework assumes three cognitive levels of human behavior: skill-based (highly practiced tasks that can be performed as more or less subconscious routines governed by stored patterns of behavior), rule-based (performance of less familiar tasks in which a person follows remembered or written rules), and knowledge-based (performance of novel actions when familiar patterns and rules cannot be applied directly, and actions follow the information processing with the inclusion of diagnosis, planning, and decision making).

This conceptual framework was adapted to the HCR method (Table 1). However, it is known from real situations and HRA practice that the distinction between skill-based and rule-based actions is often more or less arbitrary. A similar difficulty is often encountered with the distinction between rule-based and knowledge-based behavior (Reason, 1990).

3.2. Error Types
The aforementioned behavior types seem to involve different error mechanisms, which may mean radically different reliability characteristics (IAEA, 1992). Human errors are often classified into two kinds (Dougherty & Fragola, 1988):

- A slip is (a) an error in implementing a plan, decision, or intention (the plan is correct, its execution is not); or (b) an unintended action; a lapse is a type of a slip, an error in recall, for example, of a step in a task.
- A mistake is an error in establishing a course of actions, for example, an error in diagnosis, planning, or decision making.
Slips and lapses can occur during the execution of skill-based actions. A slip is, for example, an inadvertent selection of a wrong item when attempting to execute planned action. Lapses are omissions in executions of a planned sequence of actions.

Mistakes are committed when knowledge-based actions are planned. They are associated with more serious error mechanisms as they lead to incorrect understanding of the situation and to the conceiving of an inappropriate plan of resulting actions or sequences of actions. Mistakes can also occur due to an inappropriate selection of procedures or rules, based on only part of the information potentially available in order to reduce the amount of decision making required (IAEA, 1992). However, mistakes are rather less structural, and—in general—have little to do with protocol or procedure (Dougherty & Fragola, 1988).

Slips usually occur in a less time-constrained environment. Therefore, it is reasonable to believe that associated error probabilities are less time dependent. On the other hand, mistakes occur rather in a time-constrained environment and are more time dependent (it takes “process” time to think up an appropriate response in a novel situation).

It is important to be aware of different recovery potential for these types of errors. Slips and lapses can generally be corrected fairly quickly, providing there are appropriate feedback mechanisms and the plant behavior is reversible. On the other hand, mistakes are usually less easily recovered in the short term. “Mindset” problems can occur and operators can persist in attempting to implement an inappropriate plan even when faced with a lot of contradictory information. Recovery routes in such situations have to be very positive and powerful to be reliable (IAEA, 1992). A summary of the psychological varieties of unsafe acts, according to Reason (1990), is presented in Figure 2.

3.3 Incorrect Human Outputs
An early classification scheme for human errors, used in THERP (Swain & Guttmann, 1983), is related to the observable outputs resulting from incorrect human behavior. In THERP,
human errors are categorized as commissions or omissions. This classification assumes that the reliability of tasks can be assessed by decomposing them into subtasks, in a similar way as performing routine tasks can be institutionalized by writing step-by-step procedures. If a step was left out, an error of omission was said to have occurred, and if a step was performed incorrectly, an error of commission occurred (Dougherty & Fragola, 1988).

Thus, the error of commission (ECOM) is an incorrect performance of a system-required task or action, or the performance of some extraneous task or action that is not required by the system and has the potential for contributing to a system-defined failure. The error of omission (EOM) is a failure to perform a task or action (Swain & Guttmann, 1983).

This division of errors has no direct relevance to the mechanisms of their causes. For example, a person could just as easily commit an error as to omit a step because of, say, a lapse in attention. As a result, there would be no reason to expect a difference in any reliability parameter based only on this distinction (Dougherty & Fragola, 1988).

3.4. Phase of Actions

In PSA two main phases are considered: the time prior to an initiator and the time after the initiator of an abnormal event. The role of personnel, including human operators, changes over these time periods: from normal maintenance and operation activities, which are directed by familiar procedures, to incident investigation and management, and then, infrequently, to emergency response. The latter activities are driven by the event at hand and its perceived severity (Dougherty & Fragola, 1988).

Human actions and errors can be divided into three main categories, so as to relate them to the phases before and during an accident (Dougherty & Fragola, 1988; IAEA, 1992):

A. Actions and errors in planned activities, that is, the so-called preinitiator events that cause equipment or systems to be unavailable when required in an abnormal situation.

B. Errors in planned activities that lead directly—either by themselves or in combination with equipment failures—to initiating events or faults (e.g., an unplanned plant shutdown), that is, human-induced initiators.

C. Actions and errors in event-driven (off-normal) activities, that is, postinitiator events; these can be safety actions or errors that aggravate the fault sequence.

Interactions of the C category can be separated into three different types for incorporation into the PSA:

C1. Procedural safety actions.
C2. Aggravating actions and errors.
C3. Improvising recovery or repair actions.

Thus, human interactions (actions and errors) can have a variety of means of affecting the safety of the plant. It is important to include them correctly into the PSA logic structures and to understand their potential to affect risk.

3.5. Performance-Shaping Factors (PSFs)

Any factor that influences the reliability of human performance is designated as a PSF. Many factors that can potentially influence the performance of man operating a plant have been distinguished. They can be divided into three classes: (a) external PSFs, those outside the individual; (b) internal PSFs, those that can be activated within the individual himself; and (c) stressors: psychological and physiological (IAEA, 1991; Swain & Guttmann 1983).

In HRA only several PSFs are usually considered, depending on the models applied. One of the most important PSFs, and one difficult to include in the HRA models, is stress. Other important PSFs to be considered in some methods are the quality of control room design, the training quality of the staff, and other quality assurance aspects including ad-
ministrative procedures, operational procedures, and operator redundancy (Swain & Guttmann, 1983). The SLIM technique provides an analytical frame for combining PSFs to determine the overall impact on a relative basis between different tasks. An alternative is to consider only a fixed number of PSFs in a defined pattern of interaction, as it is proposed, for example, in the HCR model, developed for quantifying the operators’ nonresponse probability (Humphreys, 1988).

4. AN APPROACH TO COMPUTER-AIDED HRA IN THE CONTEXT OF PSA

4.1. Features of a Software System for Computer-Aided PSA and HRA

There is growing interest in designing supporting computer programming tools for more effectively managing the complexity of PSA and HRA and in reducing the subjectivity of assessments, as these analyses should be adequately documented for future scrutinizing and modifying when new evidence is available (“living” PSA). Lately this interest has focused on employing expert system technology (IAEA, 1990; Kosmowski et al., 1994; Poucet, 1990; Wang & Modarres, 1990). The knowledge-based systems (expert systems) are developed to imitate the knowledge of domain experts and are designed to use various information sources, including external data and expertise (IAEA, 1990).

A prototype expert system REPSAES (Reliability Evaluation and Probabilistic Safety Analysis-level 1-Expert System) has been designed (Kosmowski, Duzinkiewicz, Jackowiak, & Szczesniak, 1993). This system consists of a CAD system for graphical representation of various diagrams, databases, a shell for building an expert system, and a coordinating module, all integrated in the programming environment of MS Windows 3.1. This knowledge-based system has some interesting features, for example, the user-friendly interface with extensive graphical support using several CAD modules to represent the topological information and functional- logical knowledge. It enables effective data and knowledge acquisition as well as the iterative logic and probabilistic modeling of complex safety-related systems.

In developing logic and probabilistic models in this system, the following steps can be distinguished:

- Introducing—in a dialogue with the user—a list of basic initiating events that can initiate abnormal situations and creating classes of initiating events with their descriptions.
- Decomposing the plant into functional groups for each initiating event and constructing event trees (the construction of an event tree is partly automated when a functional logical diagram, including intersystem dependencies, was created in a CAD module, with exporting an ASCII file at the end of the edition session, which is then read by an expert system module for automatic construction of the event tree); the construction of event trees is based on a generalized object-oriented approach in which functional groups, objects, and subobjects are distinguished and dependencies among them are determined.
- The construction of fault trees for each event tree (and for each associated functional object or subobject distinguished in a given event tree) including the independent and dependent failures; the construction of a fault tree for a given functional object or subobject is partly automated if a topological diagram was created in a CAD module, with assigning attributes to “active” components and exporting an ASCII file at the end of the edition session, which is then read by an expert system module for automatic construction of the fault tree.
- Assigning to consecutive basic events of fault trees an appropriate type of the reliability model and its parameters (assigning the reliability model type can be partly automated).
- Introducing to the fault trees basic events related to human failures of Category A; that is, for situations with human performance at lower cognitive levels (this task can be partly automated), and assigning HEPs to these events.
• Modifying the event trees with regard to human-induced failure events of Category C (for situations with human performance at higher cognitive levels) and assigning HEPs to these events.
• Reduction of the fault trees by deleting basic hardware and human-induced failure events of a very low probability or human failure events of relatively low probability (comparing to basic events of independent failures).
• Sensitivity analysis and quantitative probabilistic assessment of accident sequences with determining the most important sequences.
• Considering recovery potential for most important accident sequences with introducing relevant events to related fault trees.
• Final quantitative assessment of accident sequences with a description of the consequences; assessment of uncertainty ranges in probabilistic results.

As mentioned earlier, some tasks of the analysis can be automated on the basis of the object-oriented approach, especially the construction of fault and event trees. Relevant output ASCII files contain information identifying how a given file was created. In this prototype expert system, higher priority is assigned to modeling functions performed by the expert. This means that an output file created by a module of the expert system can be modified by the human expert. Several graphical and text editors have been designed to support the user in the process of editing and modeling.

The HRA process is performed according to the already described procedures of SHARP and SHARP 1. Introducing human failure events into the structures of the fault and event trees is performed at present by the user. There is general guidance for performing these tasks (Dougherty & Fragola, 1988). Because the effects of slips during plan-directed activities usually do not propagate unassisted by other failures beyond the component(s) affected, they can be incorporated into the fault tree models. Mistakes, which are most significant in event-driven activities, because of their propagative potential, should be incorporated at the top of the fault trees or in the event trees. Recovery events are tacked at the end of the dominant accident sequences for which they apply and may not be formally included in the event or fault trees (they can be left as adjoins to cut sets). In this way, the classification effort supports the integration effort.

In the HRA/PSA screening process only more important and probable human failure events are taken into account for further analysis (Hannaman & Spurgin, 1984). Obtained HEPs are stored in a project human reliability database (Kosmowski et al., 1993). Some databases related to specific HRA methods have been designed in order to store values of attributes for the situations analyzed. Thus, the human reliability modeling process is documented.

4.2. The Selection of HRA Methods

The taxonomy of the human actions, errors, and some related factors presented in Section 3 was useful in designing a software system supporting HRA. As already mentioned, taking into account drawbacks of the conventional approach as well as the practical requirements and quality aspects of computer-aided PSA/HRA, a decision was made to design an advisory software system employing the expert system technology. This advisory HRA expert system can be employed within PSA studies, which are also based on the expert system technology, or in other safety-related analyses and decision-making activities using conventional computer PSA tools (Kosmowski et al., 1994).

After the selection of an appropriate HRA method, the assessment of HEP is carried out according to a determined procedure with regard to the attributes of the situation analyzed. Thus, the system provides a computerized framework to perform HRA, calculate HEPs, and document analyses. However, depending on the method selected, more or less external expertise from the domain expert is required in order to provide the description of the situation of interest and some factors influencing human performance.
Gradual development of the software system is based on the available—and accepted by HRA experts—techniques, which are modified when necessary for effective computer implementation. Figure 3 presents a proposal of a classification tree that facilitates the selection of an appropriate HRA technique for the analysis of human event-driven errors (Category C interactions) according to the error type and its effect. Another classification tree was proposed to select the quantification techniques for preinitiating (latent) actions and errors (Kosmowski et al., 1994).

Three levels of effort to carry out balanced PSA and HRA (Figure 3) have been distinguished: I, II, and III, which correspond to the PSA and HRA basic methodological issues (methods applied, details of modeling, and the contribution of experts required) and the relevant scopes of the computer-aided analyses (Kosmowski & Duzinkiewicz, 1993; Kosmowski et al., 1994). The development of a software system with an open architecture has been scheduled for gradual and balanced realization of the research and designing works with regard to the resources available.

In the current version of this system, a prototype version of a simplified THERP technique, based on an ASEP-HRAP approach, is available and tested. The evaluation of HEP is initiated when the human failure event is placed into the fault or event tree. Then some attributes of such an event are determined in a dialog with the user according to the procedure available in the HRA module. At the end of the session, the value of HEP is calculated and placed in a human reliability database (HRDB) related to a given PSA project.

At present, designing is focused on selected methods of scope II and III, namely, the THERP technique with graphical representation of the human reliability analysis event trees (HRAET) and a computerized version of the modified SLIM technique. The situation-specific HRAET enables representation of possible paths of human actions and failures including the

---

<table>
<thead>
<tr>
<th>Action</th>
<th>Error type</th>
<th>Effect</th>
<th>PSA/HRA effort</th>
<th>Quantification techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Slip</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mistake</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations of HRA techniques are given in Table 1. Remarks on applications of these techniques from the PSA perspective can be found in (Cacciabue, 1988; Humphreys, 1988; Kosmowski et al., 1994); CM+MS method outlined in this paper.

Figure 3. Classification of human event-driven errors and related quantifying techniques for different HRA/PSA efforts.
errors of omission or the errors of commission. The obtained structure is then assessed on the
basis of the procedure and probabilistic data available within THERP.
In the analysis of cognitive human behavior, concentrated on decision making, the construc-
tion of the operator action tree (OAT) can be useful (Hannaman & Spurgin, 1984). OAT is a
representation that identifies alternative actions on the basis of ambiguities or operator
interpretations associated with observations, diagnosis, and the selection of required responses.
The construction of OAT is also helpful in selecting an appropriate HRA technique for the
situation considered. A sample OAT is presented in Figure 4. Some modes of actions repre-
sented by sequences in Figure 4 are related to the already described classification of human
event-driven errors (Figure 3). For instance, Sequence 8 is due to an omission mistake (non-
response) and can be modeled using the HCR technique. Sequence 5 is associated with a
mistake with a potential for the commission of an error due to misdiagnosis, so the CM method
can be considered an appropriate technique for modeling such a situation. Failure mechanisms
for Sequences 2, 4, and 7 are mainly related to slips, so the THERP technique can be applied.

4.3. An Example of a Combined HRA Approach
An example of a combined HRA approach is outlined in the following. In this approach the
probabilities of the commission errors due to misdiagnosis are placed as elements in the
confusion matrix (CM). These probabilities are evaluated using the SLIM technique with
regard to several attributes assumed to be important for the situation analyzed. One of them,
considered to be very important in diagnosis, is the similarity of the plant response to a given
initiating event to plant responses to other initiating events. This approach is called a confusion
matrix approach with modeling support (Kosmowski & Duzinkiewicz, 1993), abbreviated as
CM + MS (Figure 3).

4.3.1. Remarks on quantification using the SLIM technique
The success-likelihood index method (SLIM) originates from the field of decision analysis, and
it is basically a method for quantifying the preference in a set of options (Embrey, Humphreys,
Rosa, Kirwan, & Rea, 1984). This method is used for the assessment of human reliability
assuming that human performance is affected by several PSFs that have additive effects on the

![Figure 4. An example of the operator action tree (OAT).]
probability of a successful human response. The time available to perform a task, the quality of procedure, the quality of training, and so forth, are examples of PSFs.

In order to transform the influence of these factors on human reliability, the success-likelihood index (SLI) is evaluated. It is necessary to calibrate the SLI scale for each set of tasks considered. The relationship assumed in the SLIM method is as follows:

\[ \lg p_j = a \text{SLI}_j + b \]  

where \( p_j \) is the probability of success of task \( j \), and \( a, b \) are constants. Determining the constants in Equation 1 requires at least two tasks for which the HEPs are known to be included in the SLIM session (SLIs for these tasks are also assessed). This produces two equations from which constants \( a \) and \( b \) are calculated. There are several possible sources of probabilistic data for calibration (Humphreys, 1988).

The normalized success-likelihood parameter SLI is calculated from the formula

\[ \text{SLI}_j = \sum_i w_i r_{ij} \]  

where \( \text{SLI}_j \) represents the success-likelihood index for task \( j \); \( w_i \) is the normalized importance weight in the \( i \)th PSF (the weights for all the PSF sum to 1); and \( r_{ij} \) is the scaled rating of task \( j \) on the \( i \)th PSF (the rating is scaled in terms of its distance from the ideal rating point), \( 0 \leq r_{ij} \leq 1 \) (Humphreys, 1988).

4.3.2. Using similarity measures of plant responses to support assessing elements of the confusion matrix (CM)

Mistakes, especially errors of commission due to misdiagnosis, are considered the most difficult to model and quantify. The assessed probabilities of misdiagnosis of the accident situations, in a short period after initiating events, are placed as elements of CM. They represent the probability of confusing a transient \( j \) with another transient \( k \), possibly leading to erroneous actions (Hannaman & Spurgin, 1984; Wakefield, 1988). The probabilities \( p_{jk} \) of such confusion depend on the similarity of symptoms such as alarms, enunciators, values of various process variables, or directions and rates of their changes (IAEA, 1992). In Kosmowski and Duzinkiewicz (1993), an approach to calculate measures of similarity of symptoms is described. It is easy for computer implementation and it can be then used to support the expert judgments concerning the mentioned probabilities.

Symptoms of abnormal situations are recognized in operation practice by taking into account some important variables of the process. The correct selection of observed variables is crucial for successful diagnosis. During the training process, operators acquire knowledge or personal experience of the importance of these variables. The information about transients of important process variables obtained from simulations (using a deterministic computer program for modeling the plant response for appropriate inputs relevant to the situations considered) is used to evaluate the possibility of misdiagnosis. The list of initiating events is fixed. The spectrum of plant transients is considered for the initial period of the accident progression.

The distance measure of plant responses for each pair \((j,k)\) of accident situations for the observation time \( t_p \) is calculated from the formula (Kosmowski & Duzinkiewicz, 1993):

\[ \rho_{jk}(t_p) = \frac{1}{2np} \sum_{i=1}^{n} \sum_{i=1}^{n} \alpha_i \left[ (x_{i,q}^{(i)} - x_{i,q}^{(k)})^2 + (x_{i,q}^{(i)} - x_{i,q}^{(k)})^2 \right]^{1/2} \]  

where: \( \alpha_i \) characterizes the importance of the variable \( x_i \) in the diagnosis process; \( x_{i,q}^{(i)}; x_{i,q}^{(k)} \) are components of normalized vectors for observed \( i \)th variable and its derivative, respectively, in observation points \( p \) in the observation time \( t_p \). The distance measure (dissimilarity) satisfies:
0 ≤ \( p_{j,k} \) ≤ 1. The similarity measure for each pair \((j,k)\) after the observation time \( t_p \) is calculated as:

\[
s_{p,j,k} = 1 - p_{j,k}; \quad p = 1, \ldots, r
\]

The similarity measure also satisfies: 0 ≤ \( s_{p,j,k} \) ≤ 1.

The calculated distance (dissimilarity) measures form the basis for creating distance or similarity tables for all pairs \((j,k)\) of accident situations. Depending on the obtained value of distance or similarity measure, the probability of confusion is then judged by experts, for example, using the SLIM approach. Linguistic statements concerning confusion based on the similarity measure can also be proposed; for example, high, medium, low, or insignificant, which can then form a basis for the evaluation of probability (Wakefield, 1988).

An example of a way to evaluate the influence of the dissimilarity measures on the probability of misdiagnosis when using SLIM follows. In Table 2 the results of expert opinions are presented. They were assessing normalized weights \( w_i \) for PSFs considered, and they were rating the degree of presence of PSFs for two cases: high and low dissimilarity, for which \( r_1 = 0.8 \) and \( r_2 = 0.1 \), respectively. It was also known, on the basis of two other situations assessed, that: \( SLI_3 = 0.80 \) (for which \( p_3 = 0.99 \)) and \( SLI_4 = 0.2 \) (for which \( p_4 = 0.90 \)). Based on this information, coefficients \( a = 0.069 \) and \( b = -0.060 \) were calculated from two equations, taking into account Formula 1. With these coefficients and SLIs presented in Table 2, the following probabilities of success in diagnosis were obtained (the probabilities of misdiagnosis are given in parentheses); \( p_1 = 0.953 \) \( (q_1 = 0.047) \) and \( p_2 = 0.923 \) \( (q_2 = 0.077) \). Thus, when symptoms are not similar, the probability of misdiagnosis is much lower (Case 1). This example also illustrates how sensitive the evaluations of HEPs to different expert opinions are. Thus, the quality of quantitative probabilistic assessments depends profoundly on the level of expertise and the training of the domain experts (Mosleh & Apostolakis, 1988).

5. SOME ISSUES AND CHALLENGES CONCERNING HRA IN PSA

5.1. Assessing Uncertainty

In the probabilistic evaluation of a human failure path using THERP, the intrinsic uncertainties for consecutive events are represented using the log-normal distribution, determined by the median value \( q_{0.5} \) and the so-called error factor EF (Swain & Guttmann, 1983); that is, the

<table>
<thead>
<tr>
<th>PSF _j</th>
<th>Relative Importance of PSF/10</th>
<th>Normalized Weight ( w_i )</th>
<th>Rating of Presence ( R_{n1}, R_{n2}/10 )</th>
<th>Degree of Presence ( r_{n1}, r_{n2} )</th>
<th>( SLI_1, SLI_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of information</td>
<td>10</td>
<td>0.36</td>
<td>7, 7</td>
<td>0.7, 0.7</td>
<td>0.252, 0.252</td>
</tr>
<tr>
<td>Operator training</td>
<td>5</td>
<td>0.18</td>
<td>2, 2</td>
<td>0.2, 0.2</td>
<td>0.036, 0.036</td>
</tr>
<tr>
<td>Time available</td>
<td>3</td>
<td>0.11</td>
<td>1, 1</td>
<td>0.1, 0.1</td>
<td>0.011, 0.011</td>
</tr>
<tr>
<td>Quality of procedures</td>
<td>2</td>
<td>0.07</td>
<td>5, 5</td>
<td>0.5, 0.5</td>
<td>0.035, 0.035</td>
</tr>
<tr>
<td>Dissimilarity of symptoms</td>
<td>8</td>
<td>0.28</td>
<td>8, 1</td>
<td>0.8, 0.1</td>
<td>0.224, 0.028</td>
</tr>
<tr>
<td>Totals</td>
<td>28</td>
<td>1.00</td>
<td></td>
<td></td>
<td>0.558, 0.362</td>
</tr>
</tbody>
</table>
distribution associated with ith semirandom variable is determined by two parameters: \( Q_i = (q_0.5i, \text{EF}) \). If it is assumed that consecutive failure events are independent and log-normally distributed, then the probability distribution of the compound event for this failure path (the AND operation is involved) is also log-normal and can be determined by two parameters: \( Q = (q_{0.5}, \text{EF}) \), where

\[
q_{0.5} = \prod_{i=1}^{n} q_{0.5i}, \quad \text{EF} = \exp \left\{ \sum_{i=1}^{n} (1n\text{EF}_i)^2 \right\}^{0.5}
\]

(5)

Assessing and representing uncertainties in engineering safety assessments, particularly in the case of human reliability, is usually not based on a sound theoretical framework, but rather on judgment and assumed probabilistic distributions, usually log-normal. It is obvious that using this type of distribution for representing uncertainty associated with the probability can be questioned. In engineering practice, usually more complex cases of assessments are encountered when probabilistic data are taken from various sources. Some of these data are rather hard, based on statistics, others are more or less soft, obtained by correction or fusion of information from different sources including opinions of experts. Thus, uncertainties associated with semirandom variables are represented using different distributions, so analytical probabilistic assessment of compound events is not possible in general, and therefore the Monte Carlo simulation is often employed.

Due to particularities of HRA performed in the context of PSA, an alternative framework for uncertainty representation in human reliability analysis was proposed using the possibility measures of probability and the fuzzy number calculus (Kosmowski & Duzinkiewicz, 1993). Applying the fuzzy set theory (Dubois & Prade, 1988; Zadeh, 1978, 1983) has been of interest for representing uncertain data in HRA and PSA (Sharma & Sudhakar, 1994; Tanaka, Fan, Lai, & Toguchi, 1983; Terano, Murayama & Akiyama, 1983). A framework based on fuzzy measures is also often used for the representing uncertain knowledge in expert systems (Lee, Grize, & Dehnad, 1987).

This generalized framework for data representation, combination, and handling in probabilistic safety assessments is based on the theory of evidence and the possibility theory (Kwiesielewicz & Kosmowski, 1994). There are two basic steps in the approach proposed: (a) Data with uncertainty represented by probabilistic distribution are transformed to data with possibilistic measures of uncertainty (a probability-possibility transformation); thus, all data are represented in the same form, and (b) these data are handled using the possibility theory. The logic structures (e.g., the fault and event trees) are evaluated quantitatively using the fuzzy number calculus. The theory associated with this approach is rather complex (Kwiesielewicz & Kosmowski, 1994) and is not presented here. There is, however, a considerable advantage to this approach in PSA practice, because the calculations of combined probabilistic HRA and PSA results with uncertainty ranges for assumed a-cuts are quite simple.

In most human reliability and probabilistic safety studies carried out to date, the Bayesian framework of subjective probability has been adopted and some authors are skeptical about applying some emerging theories for the uncertainty representation in PSA studies (Apostolakis, 1989; Wu, Apostolakis, & Okrent, 1990). This skepticism is associated mainly with theoretical difficulties of combining information from different sources (Dubois & Prade, 1988). The entire problem of the uncertainty representation and combining data under uncertainties, aimed at quantitative probabilistic safety assessments of complex systems, requires further research.

5.2. New Challenges in Developing HRA Methods and Supporting Computer Tools

The theory of human reliability is not yet well established and opinions have been expressed about the necessity to develop a new generation of human reliability models (Dougherty, 1990, 1993; Lydell, 1992). Human reliability can be considered as a more general problem of human factors in complex organizations (Llory, 1992). The HRA modeling approaches that now emerge
fall into four categories: procedural, temporal, influential, and contextual (Dougherty, 1993). New HRA approaches should be based on psychological aspects of human error, rather than on existing techniques (Reason, 1990). Some HRA methods mentioned in Section 2.4 are evolving in the direction of cognitive simulation environments using AI methods. Cognitive modeling is considered to be a fundamental issue for human reliability analysis (Cacciabue, 1992).

The results of simulations obtained by using these new methods for event-driven situations can convince the analyst to modify the relevant event tree or they can be included through some more important attributes (PSFs) to SLIM to support assessing the probability of failure in a given context. Such approach is proposed by authors of CES (Woods et al., 1990). It should be stressed that there is no reason to neglect all the methods used at present in engineering practice, especially those for assessing probabilities of slips for plan-driven situations (Cacciabue, 1988; Dougherty & Fragola, 1988).

It has been indicated (Dougherty, 1993) that—by examining cases of cognitive problems—analysis should remain a fundamental part of human reliability assessment. The relevant methods should include four major efforts: (a) identifying the goal matrix for the situation and any possible goal conflicts; (b) developing a functional chronology of the situation including a time matrix of the actions and decisions to be made, personnel links, and the impact of distributed decision making; (c) performing a cognitive task analysis, and (d) performing knowledge acquisition (interviewing operators, doing simulations and walk-downs). On that basis, the quantification, using, for example, the SLIM technique, will have substantial justification and the results obtained will be translatable to useful insight for further risk management.

6. CONCLUDING REMARKS
The so-called human factor significantly influences the reliability and safety of technological plants and should, therefore, be adequately included and quantified in overall probabilistic assessments. It is crucial for assuring correct results and safety-related decision making. The human reliability analysis is significantly dependent on expert opinions.

In order to facilitate the HRA and PSA studies, the use of the expert system technology is proposed, although designing a software system based on this technology requires considerable effort. According to current practice in performing PSA and HRA, three scopes of studies have been distinguished, to be covered by the supporting software system to be gradually designed. The design effort is now concentrated on scopes I and II of the software system proposed. PSA and HRA in such a knowledge-based software system will be documented to enable scrutinizing and auditing of results. Employing the expert system technology would be an important step in standardizing probabilistic studies.

In spite of critical remarks expressed recently concerning some conventional HRA methods there is no reason to neglect all the techniques used at present in engineering practice, especially those for assessing probabilities of slips. The quality of human reliability evaluation can be improved by an appropriate selection of existing conventional HRA techniques according to some attributes of the situation analyzed, including psychological aspects of human actions and errors. For that purpose, a taxonomy is proposed to be included in the advisory HRA expert system. Human reliability analysis should be not aimed only at obtaining quantitative results. No less important are qualitative results of the analyses obtained, which should be considered in improving the man–machine interactions to minimize possibilities of committing errors or to limit their consequences.

Additional research efforts related mainly to advanced HRA methods and PSA studies of scope III are required. It should—in particular—include such topics as:

- The event-driven simulation of the plant and human cognitive performance, including intention failures, aimed at modeling human reliability.
- An advanced framework for representing and treating imprecision and uncertainties in PSA and HRA at different hierarchy levels of the logic and probabilistic models of systems, including human interactions.
• Methods for combining quantitative information from various sources of different quality (credibility) including experts.
• Effective probabilistic evaluation of accident scenarios under uncertainties with regard to the equipment-oriented logic models with the inclusion of human-induced failure events and possible recovery actions.

Artificial intelligence methods, and the expert system technology in particular, offer a promising platform to deal more systematically with some challenging issues of PSA and HRA. However, the development of advanced HRA methods will require considerable research and experimental effort to verify and validate new modelling concepts.

REFERENCES


