Case Base Reasoning Approach in Industrial Accident Assessment and Management

Claudio Balducelli, Sandro Bologna, Giovanni Di Costanzo,
Adam M. Gadomski, Giordano Vicoli

ENEA - Italian National Agency for New Technology, Energy and the Environment
C.R. Casaccia, Via Anguillarese 301, 00060 Roma (Italy),), E-mail:
dicostanzo_g@casaccia.enea.it

ABSTRACT

This paper briefly describes the ENEA initiative to provide the Italian Civil Protection (ICP) with a Decision Support System in order to enhance the possibility of intervention in the field of industrial and chemical emergency. The targeted user group is the Emergency Operation Center at ICP headquarters. Our aim is to analyze the domain of intervention of ICP and provide new tools in order to enhance the capability of ICP for fast recognition of industrial disaster. We have used a system approach to integrate the new AI technologies and other products to improve fast intervention reducing negative consequences on population and environment. The CBR method has been employed to the recognition of toxic substances.

1 INTRODUCTION

Presently the most promising technologies that could give a useful contribution to the fast recognition and assessment of an industrial accident seem to be the so called Case Based Reasoning (CBR), that through a reasoning about a predefined Case Base built on the basis of experience, offers solutions to so called *ill defined problems*. That situation takes place when for lack of information available to the *decision-maker*, possible solutions even for conditions not clearly defined should be given. That's normally the situation in which people operate in real world, particularly for civil protection authorities. In that cases, even in situation not clear, but with potentially high risk, it is necessary to do some choice early, in order not to be unprepared to have to cope with situation very dangerous and in which time factor play a key role. So this research comes from a need of ICP to arrive as fast as possible to identification of an accident in industrial scope and to its precise geographical localization. Some system constraints have been identified:

- The computerized system should be only a support, not to replace the competence and skill of the personnel involved in decision-making process; it should help in a rapid and effective assessment in the range of industrial accidents.
- The system should be capable of learning and growing, while new case are inserted as new data
- As the system is targeted to inexperienced people, it should be implemented on a PC-based platform, and it ought to accept input in natural language.

The paper includes some ongoing results related to the realization of simple DSS aimed at the central control room of ICP.

2 OPERATIONAL BACKGROUND

ICP activity mainly concerns continuos monitoring of normal condition, through the control room that receives information, news and requests. Its first task is to recognize a dangerous situation that could affect the safety of people and environment, and that needs some kind of intervention for recovery. In that situation it is necessary, as fast as possible, from one hand to arrange and to start the available appropriate emergency procedures to face the actual dangerous situation, and on the other hand to get more exact information about possible consequences and development of the accident.

The DSS will give a description, of the affected area, in terms of involved people, risk object, etc. The user can see if nearby are present industries from which could be a toxic substance leakage. For that goal the DSS will use two different Databases :

- risk industries databases with information such as geographical coordinates, surrounding population, networks and infrastructures, etc.
- toxic substance database, with associated symptoms.

As new cases are entered, they will be load into the system, and the experience gained from them, will contribute to enlarge the initial accident case base. That will enhance the experience for future cases.

2.1 User specific requirements

The obtained data have been sufficient for the identification of basic user requirements and for the specification of the main functions of a managerial decision support system for the ICP emergency management role.

We have recognized the following basic user requirements :

1. Manual activation of the system on the base of registered ***symptom*** obtained from the national territory by classical communication tools , such as telephone or Fax.
2. Support in the ***recognition of toxic substances*** in the suspected zone.
3. Support in the ***identification of the industry*** which is a source of emergency, i.e. identification of place and character of incident which is a cause of emission of the toxic substances to atmosphere, to water, or to soul.
4. Support to the ***managerial intervention actions*** adequate to the recognized situation.

This paper is focused mainly on the Diagnostic module, in which case base are solved using Case Base reasoning.

3 GENERAL SYSTEM ARCHITECTURE

As shown in fig.1 the DSS has been divided in 3 different modules: Diagnostic, Predictive, and Decisional.

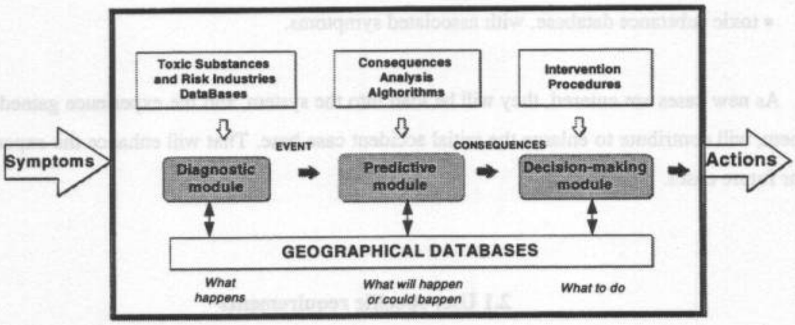

Fig. 1 DECISION SUPPORT SYSTEM ARCHITECTURE

The Diagnostic module goal is a recognition of the cause of accident. It starts from emergency symptoms acquisitions and employes the CBR method for the identification of the toxic substances and of the place of emergency source. Information about an emergency symptom are received from the different sources (police, health service, people) distributed on the national territory; operator introduces the symptom's description in natural language into computer. As first answer, the system presents as possible cause of the symptoms, industries that are in the vicinity and treat substances that may cause such symptoms. Also, other dynamic attributes of the current emergency can be hypotized on the base of the available Case Base. This last possibility is explained in the course of the paper.

The *Prediction module* goal is to indicate maximal and most probable negative consequences of the current recognized emergency state. On the user request and according the inserted data, it generates possible scenarios of the emergency which might be a consequence of the discovered accident. In the actual version this module gives as output the impact area presented on a map, obtained with an algorithm of consequences analysis. The impact area is defined as the territory in which people could be affected by the considered accident and which the operator must take under consideration during his decision-making. We need to stress that the algorithms employed to the forseen of the emergency propagation depends on many environmental and industrial factors, which, in general, enables only to assess a resonable maximal range of accidents . For instance, the propagation range of a toxic liquid and a toxic gas in the same environment conditions are usually extremely different, and requires different algorithms.

In a final version of the system it should include:

- A Knowledge Base with qualitative rules linked to different emergency states with their possible consequences.
- the numerical simulation modules which enable what-if simulation of characteristic environmental events, such as a fire and a plume propagation.

The Prediction Module concludes the Situation-Assessment function of the system.

Decisional module by using a decision-making model, suggest the operator possible actions suitable in a particular situation. On the base of current emergency situation, the Decision System should assess the level and the type of possible risk, as well as should indicate adequate tasks, using emergency instructions/procedures. This module performs an automatic transformation of the information obtained from the Diagnostics and Prediction Modules into the form of suggested actions and other data useful for a decision-making manager.

The input data are received from Diagnostic and Prediction modules. The first data is a specification of cause of registered event, such as, toxic substances and industry which is recognized as an *emergency source*. The second data is a specification of assessed possible losses-generated consequences of the recognized accident.

4 DIAGNOSTICS USING 'CASE BASE REASONING'

The approaches world-wide used to build the so called Expert Systems or Knowledge Based Systems, try to solve the knowledge base building problems utilizing *cognitive models* of involved physical domains.

The methods utilized to build such types of models are essentially based on the exploitation of *production rules* with an associated *object oriented representation* of the applicative diagnostic domain. The real possibility to exploit this type of technology often has failed against many difficulties that may raise during the phases of knowledge elicitation and formalization; in fact is often difficult to retrieve diagnostic knowledge from human experts especially in environments where it results distributed inside different sources. Consequently, the development of this type of models is difficult for the following main reasons:

- the knowledge acquisition process is difficult;
- the knowledge formalization process requires too much time and may be performed in different ways;
- the developed Knowledge Base requires frequent maintenance and updating;
- the verification and validation process of such Knowledge Base always requires the contribution of human experts;
- this type of models have low capacity of learning.

More recently new types of cognitive models were utilized in this fields: the new approaches are no more based on the capacity of *processing* and *analyzing* the situations, but on the capability of *memorization* and *retrieving similarities*. This new approach aims to experiment the emulation of the capacity of the human brain to *remember* a certain scenario and to *relate* a new scenario with an old one in terms of similarities; in fact this mechanism is regarded as the basis of human *learning* capacity(1) (2).

Some of this types of models are based on *fuzzy logic* and *neural network* techniques. Another new technology that is actually world-wide experimented and that will be considered inside this work, is the so called *Case Base Reasoning* (CBR) approach (3)(4)(5) .

The CBR approach is visualized inside fig. 2 and 3. As it is shown in fig. 2 the methods doesn't refer to a Knowledge Base but to a Case Base. The single case is represented with the following principal attributes:

1) The problem *title* that the case represents;
2) The problem *description*, using free language, that represents the principal characteristics of the case;
3) A set of *attributes* representing variables or parameters of the case itself. Such attributes can be associated to the case as a set of *questions* describing the attribute itself and whose answer furnish a value to the relevant attribute in relation with the problem described by the case. The attribute values could be *numeric* or *symbolic*.
4) A set of *suggestions* or *actions* to be executed to solve the case.

As it is visualized in fig.2, the solution method of the diagnostic problem, starts with the *characterization* process of a new case. The case is characterized introducing a general description (using natural language) of the problem and defining the attribute values (answering to questions) that are known to the system user.

Using these inputs data the system search into the case base, ordering the different cases in terms of *more or less similarity* with the configured case. The methods to find more or less similarity rely on:

1) textual comparison between cases descriptions, at level of word, set of characters and phrase, also with the help of a vocabulary of synonyms associated with the Case Base;
2) executing an attribute values analysis.

Regarding point 2, a certain attribute could be more or less *important* for the problem resolution. To the importance of the attribute may be associated a relative weight. These weights represent the so called *associative memory* that arrive to a maximum when, on the base of the *past experience*, a certain attribute value is considered determinant for the case.

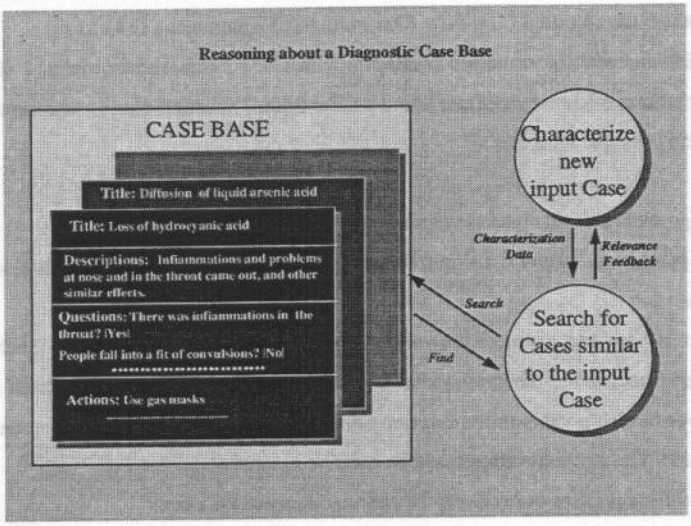

FIG.2 CBR APPROACH (1)

Some aspects of such type of functionality have similarity with the *neural processing*. In fact also a neural network works using an associative memory that is determined by the set of weights of the neural connections; it allows the discovering of similarities between different input patterns. The advantage of neural networks is that they have a completely automatic setting of the more suitable weights to recognize a certain input pattern: this automatic setting is executed during the *learning phase* and, for certain types of networks, it could be updated also during the pattern recognition phase. On the contrary in CBR the weight updating is a process that in many cases must be supervised by an expert of the problem. Anyway CBR has advantage of a more *explicit* learning process that is less dependent from mathematical algorithms and is not *visible* for the user and not sufficiently controllable by the network designer.

As it is presented in fig. 2, inside CBR process, after the searching of similar cases, using the *relevance feedback* data, the expert of the problem can evaluate the degree of goodness of the conclusion reached giving to the attributes a certain set of values. If the result doesn't fit the expectations of the expert, there are two possible hypothesis:

1) the founded case doesn't fit the expectations for a *not correct definition* of the set of weights associated to the attribute values. The lack of a good conclusion is solved updating the set of weights associated with the relative attributes.
2) the founded case doesn't fit the expectations because inside the case base there is a *lack of knowledge* relative to the considered case. That is the situation in which the actual case may be considered an *unresolved* case.

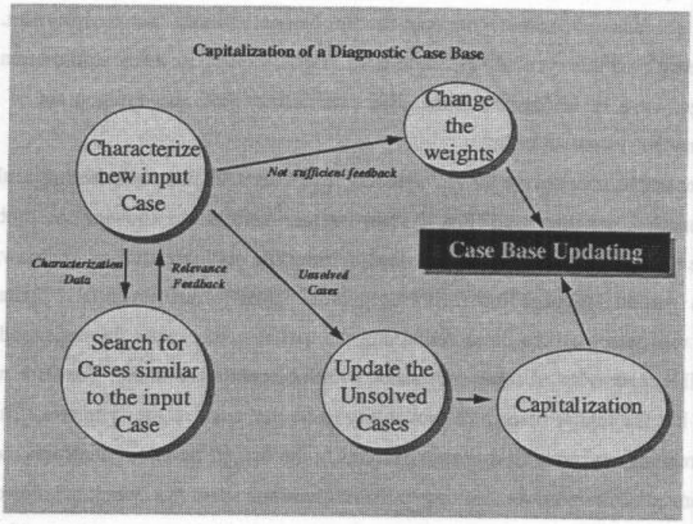

FIG.3 CBR APPROACH (2)

It is necessary to consider that unresolved cases may be generated also if the input attribute's values (provided by the system user) contain errors and contradictions that could never be present in a real case.

The more powerful CBR tools use constraints (or rules) to avoid such input data incongruences. If the case is really an unresolved case, as shown in fig. 3, this implies the saving of the case itself.

After that, by the *capitalization* process, the Case Base must be updated taking in account the founded unresolved case.

4.1 Application of CBR to the toxic substance identification.

In our application this methodology has been employed for the diagnostic process of identification of the toxic substance as possible cause of reported symptomatology. The applied software environment (4) enabled to implement Symptoms Case Base in the Diagnostic Module.

A Symptoms Case Base has been structured to allow the user to find all the toxicants that can be the cause of some symptoms for the people, animals and environment. The main characteristic of this type of case base is to allow the user to insert information in natural language, even in incomplete form. The user inserts data answering a set of predefined questions the system asks him.

For example, suppose we have a signaling involving a set of people feeling ocular troubles (lachrymation, irritation etc.) (Fig.4). Now we are interested in knowing the probable cause of these troubles. For this purpose we need to insert the only information we know, so we can type in natural language the following words: "eyes", "lachrymation", "irritation". The system will search all the cases containing the toxicants that cause the registered symptoms and will display a list of these substances with a percent of certainty. The first substance of the list has the higher percent of probability to be the cause of the symptoms. But after this first screening we surely have many toxicants in the list, all possible candidates for the above symptoms. Furthermore we also have a list of questions about the symptoms. For example we can have these type of questions: "Have you photofobia?", "Have you sight problems?", "Are there people with respiratory troubles?" and so on. These questions help the user to get more information and when the user answers these questions the system will update the list of probable toxicants with a new percent of certainty.

Some toxicants may be added to the list and some others may be eliminated. Surely when we have inserted many information about the event, we'll have a small list of cause that means a clearer indication about the probable cause of the symptoms.

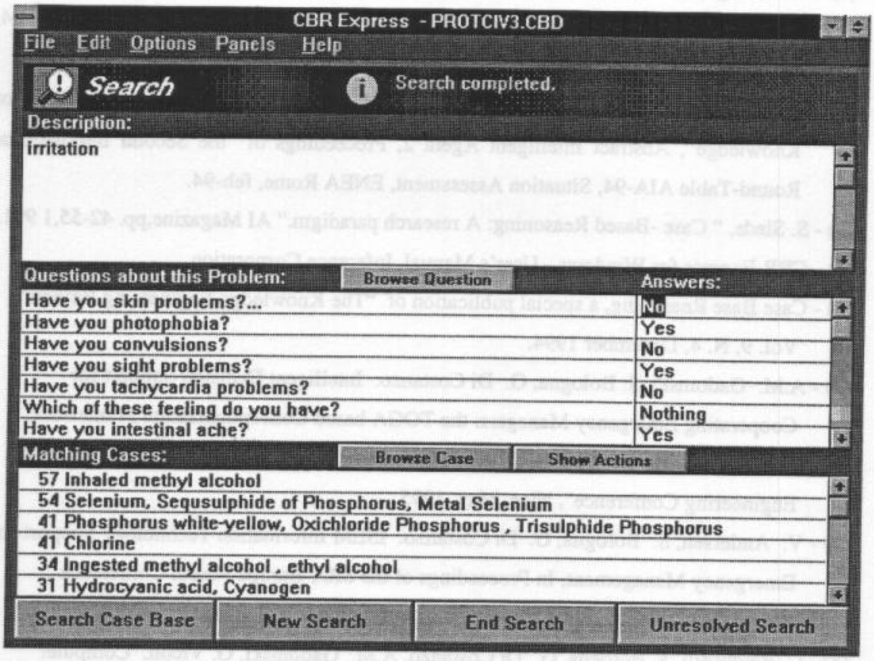

FIG.4

5 CONCLUSIONS

The presented system has been proposed as demonstrative prototype. Obtained results seem to meet the end-user requirements. Future development will concern building and validation of the knowledge base contents. Furthermore a new and more complete geographical database including information about industries in all Italian territory will be available.

However, the presented case base has to be developed in an incremental way, and the extension of the system requires an active collaboration with experts of the symptomatology related to toxic substances.

References

(1) - L. Stringa, "At The Heart of Intelligence: Memory versus Computation", Abstract Intelligent Agent 2, Proceedings of the Second International Round-Table AIA-94, Situation Assessment, ENEA Rome, feb-94.

(2) - J.M. Zytkow, "Searching for the Unknown: Machine Discoverer Assesses its State of Knowledge", Abstract Intelligent Agent 2, Proceedings of the Second International Round-Table AIA-94, Situation Assessment, ENEA Rome, feb-94.

(3) - S. Slade, " Case -Based Reasoning: A research paradigm." AI Magazine,pp. 42-55,1 991.

(4) - CBR Express for Windows - User's Manual. Inference Corporation

(5) - Case Base Reasoning, a special publication of "The Knowledge Engineering Review", Vol. 9, N. 4, December 1994.

(6) - A.M. Gadomski, S. Bologna, G. Di Costanzo. Intelligent Decision Support for Cooperating Emergency Managers: the TOGA based Conceptualization Framework. The Proceedings of "TIEMEC 1995: The International Emergency Management and Engineering Conference", Nice, May, 1995.

(7) - V. Andersen, S. Bologna, G. Di Costanzo. ISEM Information Technology Support for Emergency Management, In Proceedings of the SRA-Europe Fourth Conference , Rome, Oct., 1993.

(8) - C. Balducelli, S. Bologna, G. Di Costanzo, A.M. Gadomski, G. Vicoli. Computer Aided Training for Cooperating Emergency Managers: Some Results of MUSTER Project. The Proceedings of "MEMbrain Conference on International Aspects of Emergency Management and Environmental Technology", Oslo, June, 1995.

(9) - C. Balducelli, S. Bologna, M. Boero, G. Di Costanzo, G. Vicoli. A Computerized Support System to Cooperative Training in Emergency Scenarios Management and its Application to an Oil Port Domain. The Proceedings of "TIEMEC 1995: The International Emergency Management and Engineering Conference", Nice, May, 1995.