Journal of Hazardous Materials, 25 (1990) 93–106
Elsevier Science Publishers B.V., Amsterdam

EXPERT SYSTEM INTERFACE TO HAZARD PREDICTION MODELS

Prakash V. Rao and Phani K. Raj
Technology & Management Systems, Inc.
99 S. Bedford Street, Burlington, MA 01803

ABSTRACT

Information available at an accident scene involving the release of hazardous materials is generally insufficient and sometimes conflicting. Many times even the names of chemicals are not available to the emergency response personnel. In such cases, hazardous material behavior models alone are of little help. What is needed is a knowledge-based system which can determine with limited data from accident locations the specific chemical released (or at least the class of chemicals involved) and then exercise the appropriate hazard prediction models.

A rule based system is being developed by TMS as an add-on to the hazard assessment software system. This Expert System is expected to be used for chemical accident planning and analysis. It contains facts and rules about hazardous chemicals and their physical and chemical characteristics and possible scenarios of behavior dependent on the release and environmental conditions. The system uses this information and the responses to a set of interactively entered data to identify the chemical by eliminating chemicals which do not match the released substance. Then, behavior scenarios for the chemical are identified and applicable models are executed.

This paper describes our overall approach to the development of this expert system interface. The procedure is illustrated with a specific example.

0304-3894/90/$03.50

INTRODUCTION

In the manufacturing, handling, storing and shipping chemicals, potentially dangerous situations can occur from the accidental releases if the chemicals are hazardous. The type of danger to both public and property depends very much on the properties of the released chemical; however, the extent of danger is also a function of a number of other parameters including the release conditions, the local topography, the atmospheric conditions, and presence of other exacerbating circumstances (ignition, explosives), etc.

The literature is full of information on the remedial actions to be taken in the case of release of different types of chemicals. For example, the U.S. DOT has published the "Guide book for Initial Response to Hazardous Material Incidents:" (USDOT,1987) which provides guidance on the response action. Similar information are also available in "Emergency Handling of Hazardous Materials in Surface Transportation" (Student, 1981). The U.S. Coast Guard publication, "A condensed Guide to Chemical Hazards" (CHRIS, 1978) provides condensed information on the properties, behavior and immediate actions to take in case of chemical release. Computerized versions of chemical properties are also available in a number of databases including (to name only three) the CAMEO-RIDS and the U.S. Coast Guard's HACS and U.S.EPA's systems. The above are by no means all inclusive of the wealth of information available in the literature on the chemical properties and chemical behavior in the environment. However, all of the publications can provide guidance to the emergency response personnel as to what needs to be done **ONLY IF THE NAME OF THE CHEMICAL IS KNOWN**.

The identification of a chemical released from a fixed facility (storage tank, manufacturing plant, transfer facility, etc.) is relatively easy since the name of the chemical handled routinely in the facility will be known. The exceptions may be in facilities where a number of chemicals are used or produced as by-products of various processes. Even in such cases, the number of chemicals is likely to be small and the names of chemicals involved will be known a priori. This is in contrast to accidental releases of hazardous materials in transportation, especially when multiple tank cars are involved in an accident (as in the case of a mainline rail accident or in a rail yard). In these cases, the determination of the identity of the chemicals released becomes a very difficult task. The response action for a chemical release depends entirely on the type of chemical released and the nature of the chemical. For example, attempts to put out certain chemical fires with water (the commonly used method by most firemen) may lead to worsening of the situation because the water may react with the chemical and produce a more hazardous substance. Therefore, it is very essential to determine the name of the chemical before any emergency response action is initiated.

In the United States, the U.S. Department of Transportation regulates the movement of hazardous materials in all modes of transportation. In the case of bulk shipments of designated hazardous chemicals, placarding of the container is required. Also, in the case of shipments by truck or rail a manifest is to be carried by the truck driver and the engineer, respectively, the freight train manifest is required to list the various tank cars, their order from the locomotive, the chemical carried, etc. The chemicals are identified by a system of four digit numbers (UN or U.S. DOT number) and with color coded placards with different symbols to represent different

of chemicals. Similar requirements exist in Canada, Mexico and in European countries, though there are slight differences in the makeup of the placards and the numbering scheme. In chemical property databases, the chemicals are identified by several different schemes including, by synonyms, Chemical Abstract System (CAS) numbers, the UN/U.S. DOT number, the CHRIS three letter code, the STCC number (in rail transport way bills), etc.

If any one of these chemical identification information is available at the scene of an accident, it is a relatively easy task to obtain the properties of the chemical (either from a manual in the possession of the responder or by accessing computerized databases or by contacting chemical industry organizations such as the U.S. CMA-CHEMTREC). Unfortunately, time and again there have been accidents in which chemical releases have occurred and the emergency personnel have had no information on the consist or have had incorrect information and have, therefore, taken inappropriate action. This is because either the manifest data were lost, placards were damaged in the accident or due to human errors resulting in the wrong manifest data being provided.

There exists a need to be able to discern, in real time and with minimal instrumentation at the scene of an accident, the nature of the chemical (and if possible the name of the chemical) from visual observations of the behavior. It is the intent of this paper to indicate an approach to developing an expert system based chemical identification procedure. Such a system, it is anticipated, would be available to a responder on a laptop computer at the scene of the accident. Based on the responses provided by the operator to questions on the observed chemical behavior and other observable characteristics in the field, it is intended to identify the chemical by either class or more ambitiously provide the name of the chemical (much the same way as doctor would diagnose a patient's health condition from a series of relevant questions). Once the chemical is identified its chemical and hazard properties can be accessed from computerized or manual based databases. Hazard assessment programs (such as SAFEMODE, MicroHACS - see paper by Raj and Morris, 1990) can be run to determine the potential areas of danger, the duration of danger, and initiate appropriate evacuation or other protective actions.

The determination of the nature and type of chemical released in an accident requires certain level of knowledge and expertise (on chemical behavior). Such expertise can be found only in universities or chemical industry, neither of whose experts may be available in an emergency situation. Books and manuals are too slow and cumbersome to use in an accident response, especially if the nature of the chemical is unknown. Since speed and effective response are vital in an accident a non expert action may be insufficient or even detrimental. A computerized emergency expert system is a possible and feasible tool for assisting the response personnel. As opposed to conventional database systems, expert systems can use symbolic reasoning and therefore can deal with uncertain knowledge, incomplete data, multiple inputs, etc and provide best possible answers with explanations.

A rule based is being developed as a front end add on to a hazard assessment software system. This subsystem is called ESCAPE - Expert System for Chemical Accident Planning and Evaluation. The objective of this system is to assist the emergency response personnel in making informed guesses as to the identity of the chemicals involved in an accident. Unfortunately, because of the complexity of the

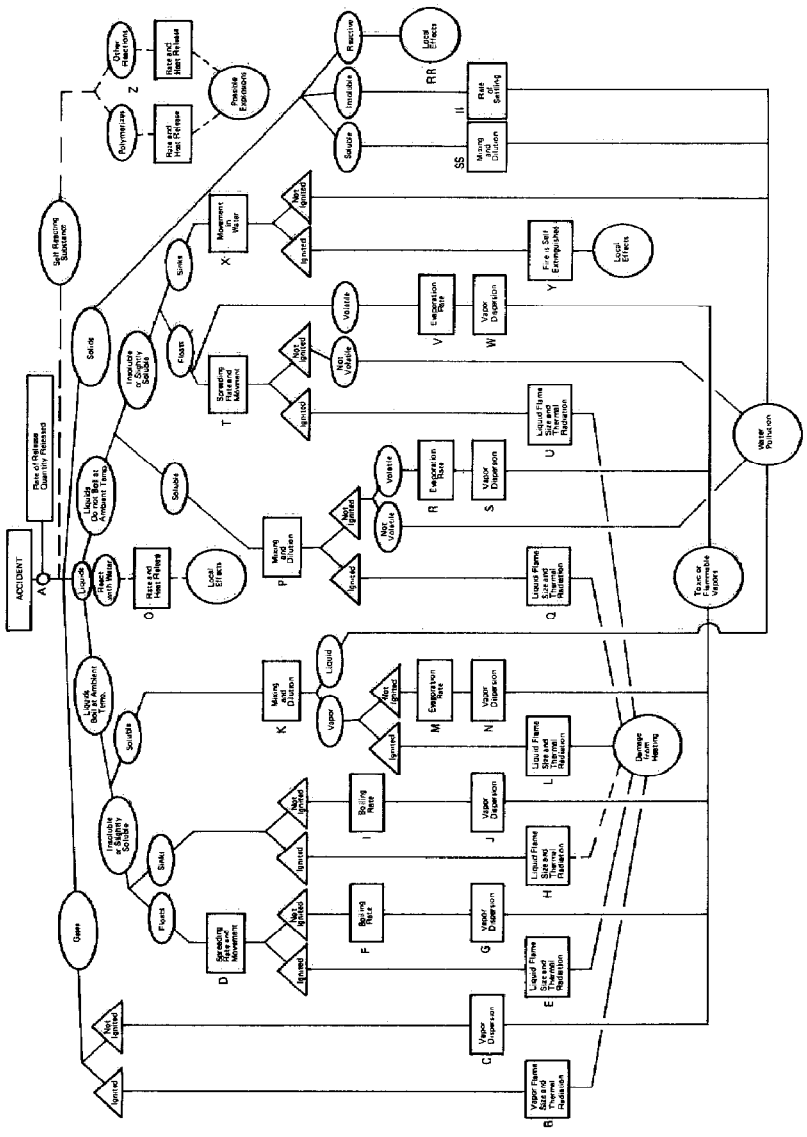

Fig. 1. Hazard assesment tree for chemical release on water (events chart; source: Raj and Kalelkar, 1974).

problems only single chemical releases are being attempted for identification. This paper provides information on the type of approach to such an identification and the design of the identification algorithm.

Design Considerations:

We consider identification of chemicals released during transportation only with the assumption that the identity of stored chemicals is more readily available. We also assume accidents involving the release of a single chemical. Extension to multiple chemicals may be a possible enhancement to the system. We approach the problem with the methods of a detective: take the evidence, look for clues and eliminate suspects. When the impossible is eliminated, the remaining, however improbable, must contain the solution.

The evidence, in this case, stems from the observable characteristics and reactions of the chemical when released into the atmosphere. We are limited to non-interactive tests since it may not be advisable to approach the chemical release point due to flames or toxic fumes. The information is limited but perhaps sufficient to identify a class of chemicals that exhibit the observed behavior.

The data available from the spill are of the following categories:

a) Definite or exactly determinable conditions. These include the substrate where the spill occurred and the atmospheric conditions at the time of the spill.

b) Properties and behavior of the chemical. These include color, apparent viscosity, volatility, nature of fumes (if any), nature of flames (if any), reaction with water, odor (if available), etc.

In one of the hazard assessment systems, the properties of the chemical and conditions at the time of the spill are used to determine the possible scenarios and therefore the applicable models. Figure 1 shows a hazard assessment for release of a substance on water. The nodes are decision points and branches are paths taken depending upon the decision. The ovals represent the nature of the substance (the basis for the decision and branch), the boxes suggest the parameters to be obtained for hazard assessment, and the triangles select between the ignited and not-ignited branches. Following a particular path we have the effect of the accident and the model to apply. For example, a liquid which does not boil at ambient temperature, is insoluble in water and does not evaporate leads by the path AT to the effect "Water Pollution".

We attempt to use the observable behavior and the atmospheric conditions to determine the properties of the chemical, and therefore the chemical. Observable behavior is dependant upon the chemical properties and the atmospheric condition. We use this dependance in the reverse direction to eliminate chemicals which do not exhibit the behavior. For example, if there is a fire, chemicals which are not combustible or flammable can be eliminated. A point to note is that while an observation eliminates a class which does not exhibit the behavior, absence of the behavior does not necessarily eliminate anything: for example, absence of a fire does

not imply that the chemical is not combustible. Thus by eliminating impossible cases, the search space is reduced.

This paradigm - eliminate the impossible - attempts to discover all that can be deduced from the given set of facts. By this method, when all facts have been considered, if there still is not enough information to fix upon a definite solution or set of solutions, the set of partial matches are "possible" solutions. The method assumes a finite solution space. The search gets faster as the search space is reduced.

To account for uncertainty, only definite mismatches are eliminated. Possible mismatches are maintained. To distinguish between "definite" and "possible" matches, certainty factors are used. This can be interpreted as a "likelihood factor", and ranges from 0 to 1. When a chemical is "impossible" it is given the likelihood factor of 0. This controls the search -chemicals are tried by order of likelihood. Initially, all chemicals are equally likely. As more evidence is collected, the likelihood factors are changed based on the certainty factor of the evidence and the probability that the chemical will exhibit the evidence. If the likelihood factor is 0, the chemical is eliminated from the list of possible chemicals. When more likely cases have been tried and eliminated, less likely cases are still tried.

Implementation:

In most expert systems, indeed, in most AI systems, the method of finding a solution is by searching for the goal state in a solution space. The search may be conducted by starting at the initial state and moving to successive states (sub-goals) based on the knowledge available until the goal is reached. This is called Forward Chaining. The search may alternatively be conducted by taking one of the possible goal states as a hypothesis and proving that it can be reached from the initial state. This is called Backward Chaining. Combinations of these are also used to make the search more efficient.

Two kinds of problems generally addressed by expert systems developers are synthesis and diagnosis. Forward chaining is generally used for solving synthesis problems - starting with known "ingredients" and working them to the solution. This is particularly applicable when there are fewer starting parameters and too many possible goals. To establish or reject each of these solutions is not feasible. Backward chaining is generally used for solving diagnosis problems - establishing that the symptoms match with the conditions known of the suspected goal state.

Our problem is one of identification. It is similar in principle to diagnosis problems, but our approach is different from those of standard diagnosis systems. We do not attempt to match the chemical. Rather, we eliminate mismatches and narrow down the search space towards the solution. This is due to the peculiar nature of our problem. True identification of a chemical can be achieved by studying the results of tests that demonstrate its chemical properties. A medical diagnosis system, MYCIN (Shortliffe, 1976) uses symptoms and results of clinical tests to identify possible infections and suggest remedies. We are restricted to observable properties and behavior - "symptoms" - and, therefore, it is more practical to eliminate definite mismatches than to attempt to establish definite matches with limited information.

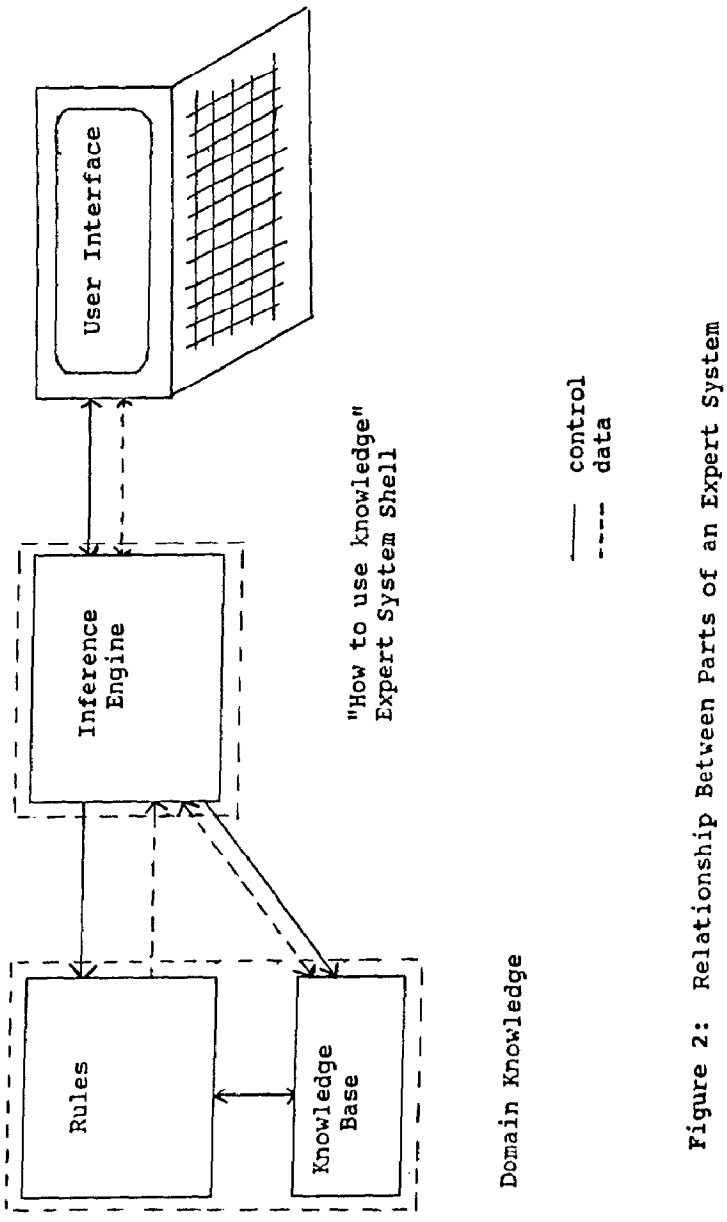

Figure 2: Relationship Between Parts of an Expert System

It is believed that all logical decisions taken by human beings result from application of heuristic rules (if-then constructs) to the current situation. Rules are useful for capturing elementary logic. Knowledge can be represented in the form of rules.

An expert is one who has two kinds of knowledge:

1) Knowledge about the domain, and

2) Knowledge about how to use knowledge.

The former is specific to the domain, the latter is necessary in all domains. The former is captured as the rules in a rule-based system. The latter is the technique used in the application of these rules, including the search mechanism (forward- or backward-chaining), the pattern matching algorithm and the conflict resolution strategy.

In addition to rules and rule application mechanism (the Inference Engine), an expert system has domain specific knowledge to which these rules are applied.

Figure 2 shows the relationship between the logical parts of an expert system. The inference engine controls the application of the rules upon the knowledge, and the modification of the working knowledge depending upon the rules applied. The inference engine also interacts with the user to obtain information required for application of certain rules, and to present to the user the solution obtained.

An important issue in the design of an expert system is the right organization of knowledge. The objective of this knowledge representation is to model the real world by abstracting relevant attributes of the pertinent objects in the real world.

The objects of interest to us are:

1) Chemicals (in particular, liquids that are generally transported) and their properties.

2) Conditions under which spill occurred

3) Behavior of the spilled chemical under the conditions above.

In addition to these objects, certain control elements are modeled to keep track of the operation of the expert system.

The chemical properties that are not observable are of no consequence to us. Only physical properties and observable reactions of the chemicals are modeled. Relevant atmospheric conditions include (approximate) temperature, humidity, wind speed, sun/rain/snow and conditions on land or water depending on where the spill occurred. Behavior of the chemical is captured in elements which state whether the property **p** is present or unknown. Any form of classification serves to make to representation more structured. Since chemical properties are not considered, observable behavior is used to classify the chemicals. The classification is hierarchical. Frames are used to represent the hierarchy. Figure 3 shows a partial structure of the representation.

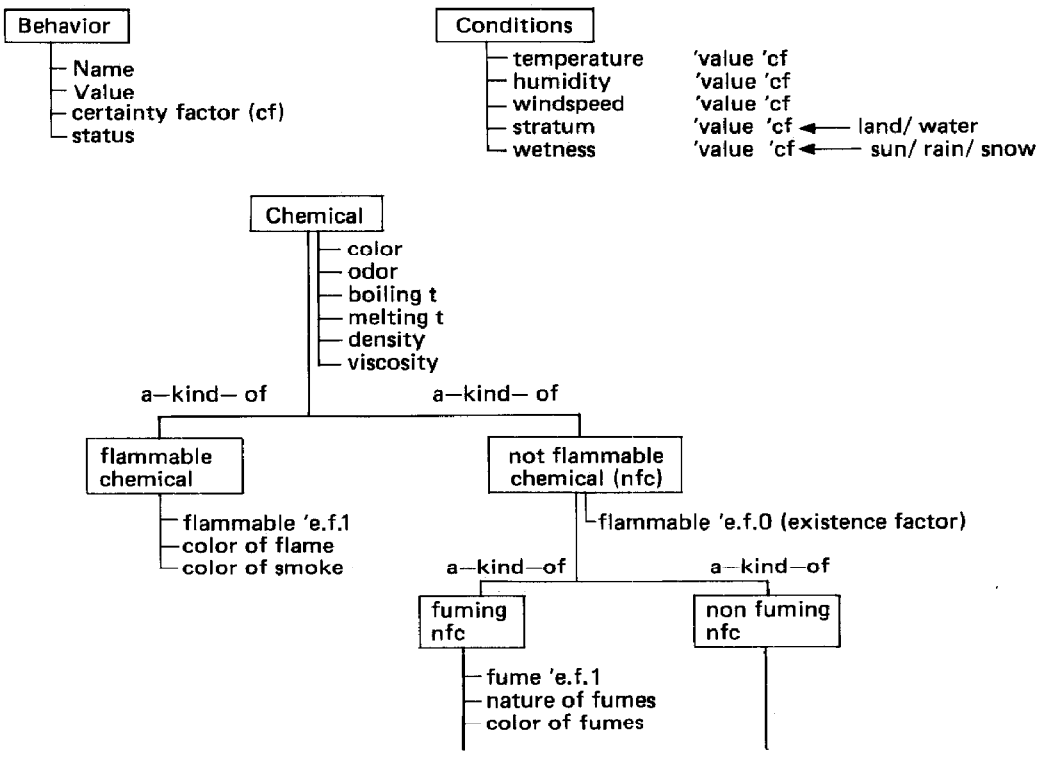

Figure 3: Partial Structure of Knowledge Elements

All chemicals in a class have the same set of properties. The parent classes have a subset of properties and the children classes have a superset [Specialization down the hierarchy]. The properties of the parent class are inherited by the child class.

The properties are classified under physical and observable properties. Observable properties include combustibility, volatility, color, color of fumes or flames as applicable. Physical properties include boiling temperature, melting temperature, toxicity, etc. Associated with each observable property is an existence factor (e.f.) in the range 0 to 1 which is the probability that the chemical exhibits the property. This factor may be a function of temperature, humidity, etc.
The functioning of the system is aimed as follows:

Initially, a working copy of the chemical information is made. All likelihood factors are set to 0.5 [When nothing is known, every possibility is equally likely]. All the observable property slots are given the value 'blank. As more information is obtained, these values are changed to 'yes if known, or 'unknown.

The system moves to the final phase when all property values are non-blank and all behavior elements are inactive.

Until then, the execution is as follows:

If there are no active behavior elements, from the list of possibilities sorted on likelihood factors, one of the most likely elements is selected. If the chemical has a property **p** whose value is blank and the parent of the chemical does not have a slot for the property, a behavior element is created for **p**. The user is asked for the existence of the behavior, and the confidence (range 0 to 1) with which the existence is stated, or 'unknown if the information is unknown. The behavior element is made active. The certainty factor (c.f.) is set to the confidence in the user's response.

While a behavior element **p** is active, if there is a chemical **c** with a property slot for **p** set to 'blank, the property slot is set to value of the behavior element. If the behavior is not unknown, the likelihood factor of the chemical is recalculated as follows:

If the absence of the observation does not signify anything, e.g. absence of a fire does not signify that the chemical is not combustible, if the behavior is not observed, likelihood factors are left alone.

If **p** is 'unknown, l.f. is left alone.

t = (c.f. of behavior **p**) * (probability that **c** exhibits the behavior)

If t = 0, l.f. = 0
else l.f. = average of old l.f. and t

If there are no more chemicals with property slot for **p** set to 'blank, the behavior element **p** is set inactive. If **p** is not 'unknown, the possibility list is sorted on l.f. If an l.f. is 0, the chemical and its descendants are eliminated from the list of possibilities.

In the final phase, if no possibility or set of possibilities can be confirmed, the list of possibilities is presented to the user.

An Example:

Consider the identification of LNG with information of an accidental spill over water. LNG is generally transported over water in a cryogenic form. If the chemical is not ignited, it boils rapidly and produces dense white clouds by condensing the water vapor around it. If the chemical is ignited, it burns violently with bright yellow flame and no smoke.

The database initially contains the following chemicals:

1) Gas

- Ammonia - flammable 0.0, forms toxic or flammable vapors 0.8;
- Chlorine - flammable 0.0, forms toxic vapors 0.8;
- LPG - flammable 1.0, forms toxic/flammable vapors 1.0;
- LNG - flammable 1.0, forms toxic/flammable vapors 1.0.

2) Liquid

Liquids that boil at ambient temperature:

- Ammonia - reacts with water 0.8, soluble in water 1.0,
- Chlorine - sinks 1.0, reacts with water 0.0;
- LPG - flammable 1.0;
- LNG - flammable 1.0;

Liquids that do not boil at ambient temperature:

- Ethyl Alcohol - soluble 1.0, flammable 1.0, volatile 0.8;
- Sulphuric Acid - soluble 1.0, reacts with water 1.0;
- Nitric Acid - soluble 1.0, flammable 0.0, volatile 0.0;
- Kerosene - soluble 0.0, flammable 0.9;
- Diesel - soluble 0.0, flammable 0.7, fume 0.0;
- MEK - soluble 1.0, volatile 1.0, flammable 1.0;
- Phenol - soluble 1.0, flammable 1.0

3) Solid

- Aluminum Sulphate - soluble

Note the duplicate entries for gases under liquids that boil at ambient temperature. All gases are potentially transportable as liquids, either pressurized or cryogenic (exceptions: substances like oxygen which cannot be liquified under pressure). The rule base includes rules of the following types:

If the substance boils at ambient temperature c.f.1.0

——→

substance fumes c.f.1.0.

If the substance reacts with water c.f.1.0

——→

substance fumes c.f.0.5.

Also, some chemicals have the quality "fumes" explicitly stated. Heavy oils such as diesel do not fume and are so noted (diesel 'fumes 'c.f.0.0). The color of the fumes of chlorine (when the observer is close enough to observe the color) is greenish-yellow. The color of the flames of some chemicals are also known (LNG is bright yellow, LPG is bright red, etc.). Some chemicals burn without smoke (LPG, LNG, ethyl alcohol) while others give off thick smoke (diesel).

In this example, the exact order of queries that the system asks is not predictable at this stage. Assuming that the order is immaterial, we have the following scenario.

The initial likelihood factors are all 0.5. All gases and solids are eliminated since a liquid has been spilled. Assuming the chemical is not ignited, a behavior element for "fumes" is created and set to active. The value of this behavior element is 1.0 since fumes are observed.

The new likelihood factors are calculated. If it is not known if the chemical fumes or not, the l.f. is the old l.f. (t = old l.f.). The following is the new state.

Ammonia both boils at ambient temperature and reacts with water. The likelihood of ammonia fuming is 1.0, from the rules.

t = (c.f. of behavior fumes) * (probability that ammonia fumes).

= 1.0 * 1.0 = 1.0

l.f. = average of old l.f. and t = $\frac{0.5 + 1.0}{2}$ = 0.75

Similarly, l.f. of other chemicals are obtained:

Chlorine = 0.75
LPG = 0.75
LNG = 0.75
Diesel = 0.0

Everything else remains at 0.5. Diesel is eliminated. While others are not unlikely, ammonia, chlorine, LPG, and LNG are more likely. With more information, more chemicals can be eliminated and fewer chemicals emerge in the "more likely" set.

If the substance is ignited, chemicals which are not flammable are eliminated. From the color of the flame and the fact that LNG burns without smoke, LNG is identified.

In Conclusion

The system is proposed to be developed on an IBM PC in LISP. The search mechanism desired is forward chaining, the pattern matching will be performed by a Rete-like network and the conflict resolution strategy Lexical.

While the system can be made to confirm to the designed behavior, changes will be incorporated if they prove more efficient. Extensive tests need to be run before a consistent behavior can be established.

Possible extensions to this design are:

1) Ability to handle spills involving more than one chemical.

2) Generalization from liquid spills to release of any chemical.

3) Ability to consider unkowns with default reasoning, rather than ignore them as we have done

REFERENCES

Barr et al.,1982, "The Handbook of Artificial Intelligence, Vol 2", Barr,A., Fiegenbaum,E.A., William Kaufmann, Inc., Los Altos, CA.

Barr et al.,1989, "The Handbook of Artificial Intelligence, Vol 4", Barr,A., Cohen,P.R., Fiegenbaum,E.A., Addison-Wesley, Reading, MA.

CHRIS,1978, "A Condensed Guide to Chemical Hazards",CG-446-1, Vol 1 of Chemical Hazard Response Information System, US Coast Guard, Washington, DC.

Hayes-Roth et al.,1983, "Building Expert Systems", Hayes-Roth,F., Waterman,D.A., Lenat,D.B. (Eds.), Addison-Wesley, Reading, MA.

Pederson,1989 "Expert System Programming", Pederson,K., John Wiley & Sons, Inc., New York, NY

Raj, P.K. and A.S. Kalelkar, (1974), "Assessment Models in Support of the Hazard Assessment Handbook (CG-446-3)", NTIS #AD-776617, Technical report to the U.S. Coast Guard, Washington, DC 20590, January 1974.

Raj & Morris,1990, "Computerized Spill Hazard Evaluation Models", Raj,P.K., Morris,J.A., Paper to be presented at the 2nd International Special Meeting on "Major Hazards in Transportation and Storage of Pressure Liquified Gases", May 1990, Queens University, Ontario, Canada.

Shortliffe,1976, "Computer-based medical consultations: MYCIN", Shortliffe,E.H., American Elsevier, New York, NY.

Student,1981, Emergency Handling of Hazardous Materials in Surface Transportation", Student,P.J., Bureau of Explosives, Association of American Railroads, Washington, DC.

USDOT,1987, "Guide book for Initial Response to Hazardous Material Incidents", ERG (DOT P 5800.4), US Department of Transportation, Washington, DC.

Waterman,1986, "A Guide to Expert Systems", Waterman,D.A., Addison-Wesley, Reading, MA.

Winston,1984, "Artificial Intelligence", Second Edition, Winston,P.H., Addison-Wesley, Reading, MA.