Computers and Chemical Engineering 33 (2009) 371–378

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Learning HAZOP expert system by case-based reasoning and ontology

Jinsong Zhao[a,*], Lin Cui[b], Lihua Zhao[b], Tong Qiu[a], Bingzhen Chen[a]

[a] *Department of Chemical Engineering, Tsinghua University, Beijing, 100084, China*
[b] *College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 100029, China*

ARTICLE INFO

Article history:
Received 1 November 2007
Received in revised form
13 September 2008
Accepted 4 October 2008
Available online 22 October 2008

Keywords:
HAZOP
Case-based reasoning
Ontology
Process safety

ABSTRACT

To improve the learning capability of HAZOP expert systems, a new learning HAZOP expert system called PetroHAZOP has been developed based on the integration of case-based reasoning (CBR) and ontology that can help automate "non-routine" HAZOP analysis. PetroHAZOP consists of four modules including case base module, CBR engine module, knowledge maintenance module and user graphical interface module. Within the case base, HAZOP analysis knowledge is represented as cases which are organized with a hierarchical structure. Similarity-based case retrieval algorithm is also depicted to find the closest-matching cases. In order to enhance the case retrieval, a new set of ontologies for CBR-based HAZOP analysis is created by integration of existing ontologies reported in literature. Finally the application of PetroHAZOP is demonstrated by two case studies of industrial processes.

1. Introduction

Safety is an important issue in process design and operation in the chemical process industry (CPI). It is even more critical for modern chemical manufacturing processes which are often operated under extreme conditions to achieve maximum economic profit, or have to undergo changes of customer demands. The importance of safety analysis in process operation is well recognized after occurrence of several tragic accidents that could have been avoided if adequate process safety analysis had been done. To ensure safe operation, process hazard analysis (PHA) is very important to proactively identify the potential safety problems and recommend feasible mitigation actions. Among the available PHA techniques, hazard and operability (HAZOP) analysis is the most widely used one in the CPI. HAZOP analysis done by human teams, however, has the following shortcomings: time consuming, laborious, expensive and inconsistent. To solve these problems, various model and/or rule-based HAZOP expert systems have been developed during the last decade, which was respectively reviewed by Venkatasubramanian, Zhao, and Viswanathan (2000) and McCoy et al. (1999). These systems, however, can only address "routine" or process-generic HAZOP analysis. "Routine" HAZOP analysis means that its reasoning logic can be applied to different processes while the "non-routine" HAZOP analysis means that its reasoning logic is process specific or plant specific. Generally analysis of deviations generated by using guidewords "other than", "as well as" and "part of" are "non-routine". As a result, these kinds of deviations are hardly addressed in literature about HAZOP expert systems.

In the CPI, "routine" HAZOP analysis roughly occupies 60–80% while "non-routine" HAZOP analysis occupies 20–40%. Due to the lack of self-learning capability in existing HAZOP experts systems, the knowledge of "non-routine analysis" can be hardly formulized and reused for similar chemical processes, and the "non-routine" HAZOP analysis still needs to be addressed by human experts.

To evaluate the output quality of the signed directed graph (SDG) model based HAZOP expert system HAZID developed by McCoy et al. (1999), five industrial plant systems which had not been used during the model development stage were selected as a test set (McCoy, Wakeman, Larkin, Chung, & Rushton, 2000). The output of HAZID was compared against the results of conventional HAZOP study which was done by human teams. Table 1 shows the test results which are quite interesting. According to Table 1, the percentage of scenarios identified by conventional HAZOP also identified by HAZID ranged from 33% to 60%, and the percentage of scenarios identified by HAZID which were judged to be corrected ranged from 33% to 83%. Moreover, the percentage of scenarios identified by HAZID which were judged to be correct and of interest was much lower, ranging from 9.5% to 29%. In other words, severe completeness and correctness issues existed in HAZID. The unfavorable performance of HAZID was attributed to a few factors such as the quality of the unit model, lack of fluid property data, the plant complexity and human judgment variance. In fact, there is another more important factor that was unrecognized and that factor is knowledge representation limitation. The "non-routine"

* Corresponding author. Tel.: +86 10 62783109.
E-mail address: jinsongzhao@tsinghua.edu.cn (J. Zhao).

0098-1354/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2008.10.006

Table 1
Summary analysis of trial results of HAZID (McCoy et al., 2000).

	Plant systems				
	Absorber system	β-trichloroethane system	Propane rectification system	Benzene storage	Separation system
Scenarios identified by conventional HAZOP also identified by HAZID (%)	36	33	60	50	53
Scenarios identified by HAZID which were judged to be correct (%)	49	33	69	83	53
Scenarios identified by HAZID which were judged to be correct and of interest (%)	9.5	29	24	27	N/A
Protections identified by HAZID which were judged to be correct (%)	9.5	29	N/A	77	N/A

HAZOP analysis could not be represented by the existing models. Therefore, it is clear that there is much room for improvement in knowledge representation for HAZOP expert systems.

It is worth noting that consistence and completeness are critical in HAZOP analysis because neglect of any potential hazard may even result in disasters. Investigation results of past industrial accidents, e.g. the tragic BP Texas city plant accident occurred in March 2005, have proved that poor quality of PHA is a major root cause of accidents occurred in the CPI.

Recently case-based reasoning (CBR) technology (Kolodner, 1993; Aha, 1998) has been integrated into HAZOP automation technology by researchers at Purdue University to enhance the self-learning capability of HAZOP expert systems (Zhao, Bhushan, & Venkatasubramanian, 2005). However, the case-based reasoning they proposed aimed to facilitate modification of the existing models and creation of new models based on the knowledge in the existing models. The "non-routine" HAZOP analysis still relied on the human team. To solve this problem, a new learning HAZOP expert system called PetroHAZOP has been developed in this paper based on the integration of CBR and ontology that can help automate both "routine" and "non-routine" HAZOP analysis.

This article is organized as follows. In Section 2, HAZOP analysis and related work on the automation of HAZOP are briefly described. In Section 3, the integrated methodology for HAZOP analysis is explained. Section 4 contains two industrial application examples that illustrate how the proposed HAZOP expert system can help with improvement of "routine" and "non-routine" analysis. Finally, contributions of this work are summarized and discussed in Section 5.

2. HAZOP

HAZOP was firstly introduced by ICI (Imperial Chemical Industries, UK) for identifying hazards in chemical plants in 1960s. HAZOP study is accomplished by a HAZOP team through a collective brainstorming effort that stimulates creativity and brings about new ideas of the potential hazards including their cause–effect relationships.

Generally the chemical process is divided into sessions called "analysis nodes" before study. Then meaningful deviations in every analysis nodes are generated by combining process parameters and HAZOP guidewords including MORE OF, LESS OF, NONE, REVERSE, PART OF, AS WELL AS and OTHER THAN. For each deviation, the HAZOP team has to identify all of its credible causes and all of possible adverse consequences. Once the causes and consequences are recorded, the team has to list the existing safeguards for the identified hazards and give necessary recommendations accordingly for hazard mitigation if the required risk level cannot be achieved by the safeguards. The process is repeated deviation by deviation and node by node until the analysis of the whole process is completed. The conventional HAZOP study procedure is presented in Fig. 1.

To complete the HAZOP analysis of a typical chemical process, it takes about 1–8 weeks for a HAZOP team with 4–8 members. It is widely accepted that HAZOP analysis is an extremely time consuming and effort consuming process. An estimated including direct and indirect costs $5 billion is spent annually by the CPI to perform PHAs and related activities. The estimated cost of process hazards reviews is about 1% of sales or about 10% of profits for a big chemical company. Moreover, the quality of HAZOP analysis depends on the knowledge and experience of the HAZOP team. Therefore, incompleteness and inconsistence usually are the drawbacks with regards to HAZOP done by human teams. Given the enormous amounts of time, effort and money involved in performing HAZOP, there exists considerable incentive to develop intelligent systems for automating the process hazards analysis of chemical process plants. An intelligent system can reduce the time, effort and expense involved

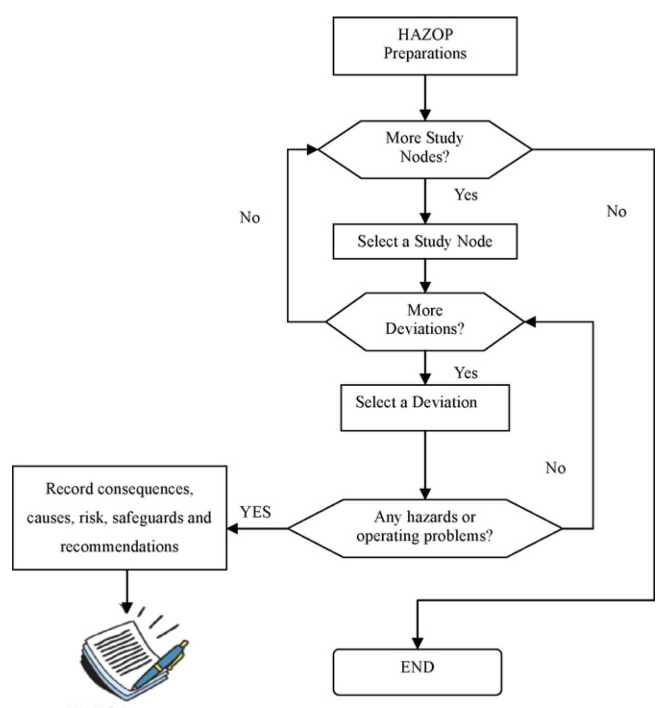


Fig. 1. HAZOP study procedure.

in a HAZOP, make the analysis more thorough, detailed, and consistent, minimize human errors, and free the team to concentrate on the more complex aspects of the analysis which are unique and difficult to automate (Venkatasubramanian et al., 2000). The HAZOP analysis that is difficult to automate generally refers to "non-routine" analysis discussed in the above section. In what follows, case-based reasoning and ontology are integrated for the automation of HAZOP analysis that was considered difficult to automate before by using the traditional model-based or rule-based approaches.

3. HAZOP expert system by the integration of CBR and ontology

3.1. Case-based reasoning (CBR)

Experts often find it easier to relate stories about past cases than to formulate rules. Similarly it is true in the HAZOP analysis domain that rules or models are hard to construct to automate "non-routine" analysis. To overcome this problem, an important artificial intelligence technique – CBR is adopted to augment the reasoning machines embedded in the existing HAZOP expert systems. CBR is both a pattern for computer-aided problem solvers and a model of human cognition. The central idea is that the problem solver reuses the solution from past cases to solve a new problem. In this way, valuable experiences that are difficult to formulate into rules or models could be utilized for solving new problems.

In a CBR system, the problem solving process includes four phases (Aamodt and Plaza, 1994), namely 4R's: Retrieve, Reuse, Revise and Retain as shown in Fig. 2. Basically, knowledge and experience are stored in the form of cases in a case base. The content of a case is made up of three parts: the problem/situation description, the solution, and the outcome. The outcome is not needed but could be added to suggest solutions that work and use cases with failed solutions to warn of potential failures. When a new problem is submitted, CBR system retrieves the similar cases from the case base using certain similarity algorithm based on the predefined indexes. Indexes should be abstract enough to retrieve a relevant case in a variety of future situations, and also should be concrete enough to be easily recognizable in future situations. Then the solutions of retrieved cases are adapted to solve the new problem if necessarily. Finally, the new problem description and its solutions are retained as a new case in the case base for future use.

Most of the CBR applications do not go through all the above four phases (López-Arévalo, Bañares-Alcántara, Aldea, Rodríguez-Martínez, & Jiménez, 2007). In this paper, the adaptation is done by the users because it is highly domain dependent and requires verification of the solution performance.

3.2. Ontology

Human experts are indispensable in HAZOP analysis of any chemical processes even though various expert systems can be designed to facilitate the process. Different experts especially from different organizations have different jargons with regards to the descriptions of the analysis objects and results including causes and consequences of hazards. That is to say, there is no standard to represent the HAZOP analysis domain information. This increases the difficulty of CBR for different users. To settle the terminological and conceptual incompatibility problem, a new set of ontologies for CBR-based HAZOP analysis (CHA) is created in this paper by integration of existing ontologies reported in literatures.

"Ontology" is a term of philosophy originally, which refers to the subject of existence. Artificial intelligence (AI) borrows the old term from philosophy and gives it new wonderful meanings. In AI, there are a number of definitions of ontology. However, the definition given by Gruber is accepted by majority of researchers: an ontology is an explicit specification of a conceptualization (Gruber, 1993).

HAZOP analysis of chemical process needs knowledge from areas such as chemistry, chemical process engineering, safety engineering, electrical engineering and so on. A large-scale ontology OntoCape was constructed by the research group of Professor

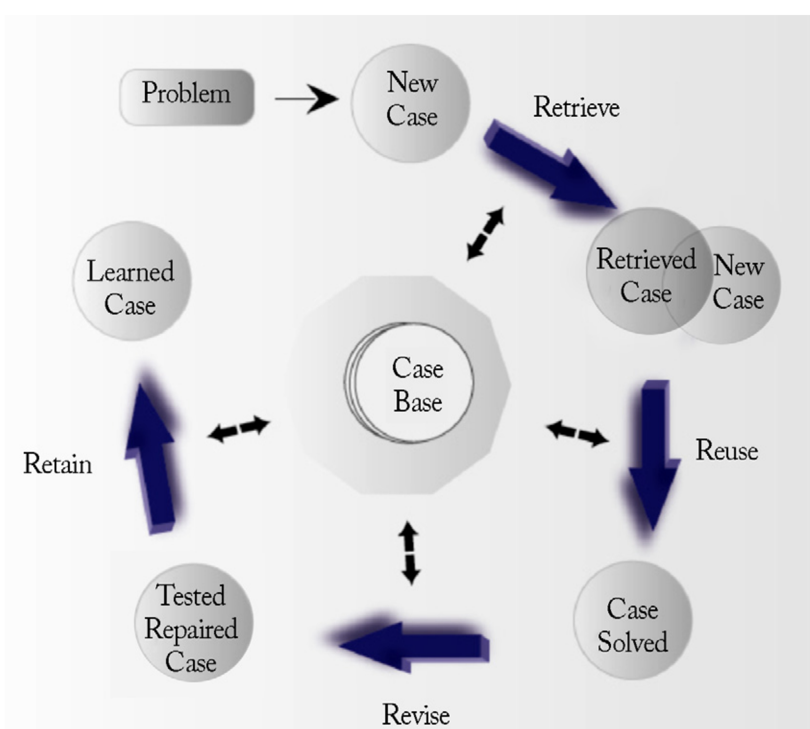


Fig. 2. Problem solving phases with CBR.

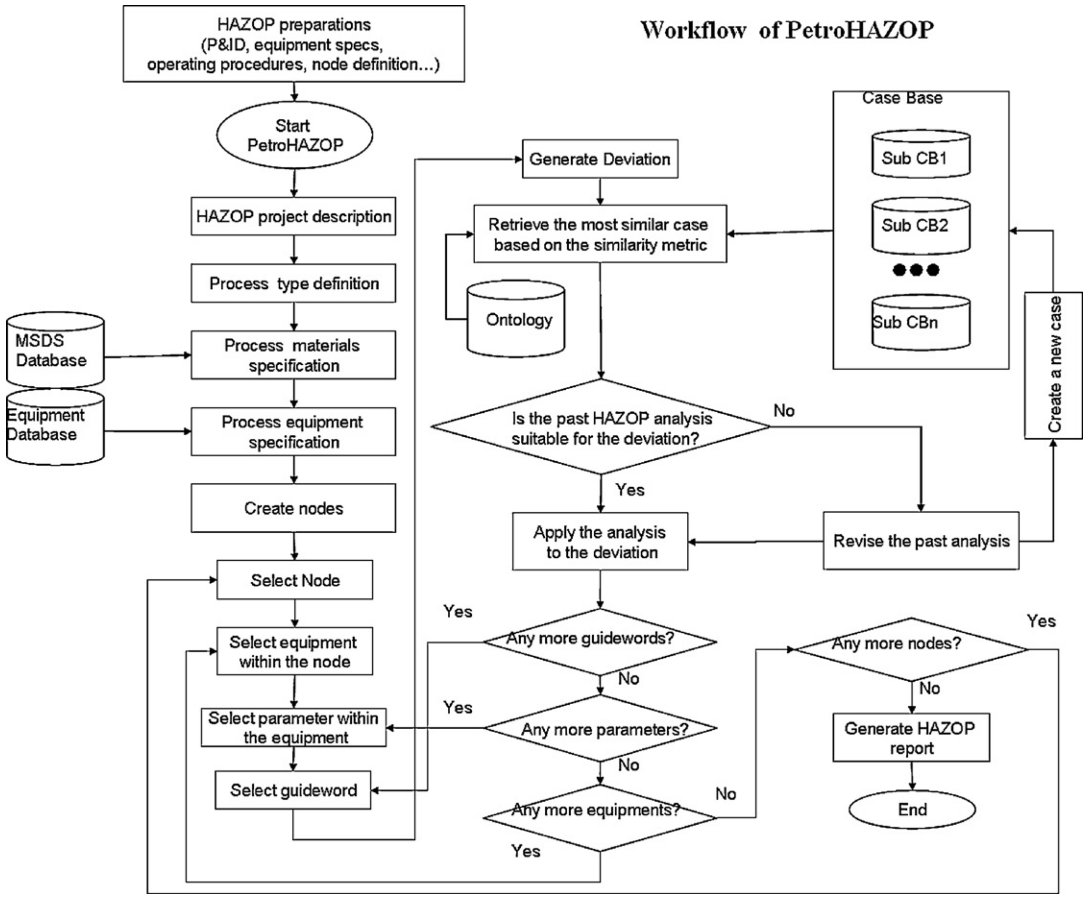


Fig. 3. Workflow of PetroHAZOP.

Marquardt for chemical process engineering (Morbach, Yang, & Marquardt, 2007). To facilitate the information sharing among the HAZOP analysis expert systems developed by the laboratory of Professor Venkatasubramanian at Purdue University (Venkatasubramanian et al., 2000), process simulation packages such as Aspentech's BatchPlus and documentation tool such as Dyadym's PHAPro, operational related ontologies and safety related ontologies were created (Zhao, Bhushan, & Venkatasubramanian, 2003). Batres et al. created an upper level ontology based on ISO 15926 that had already been used for knowledge queries in HAZOP (Batres et al., 2007). Based on the major concepts and ideas from the above ontologies, six ontologies are created for case-based HAZOP analysis. They respectively are process ontology, process unit ontology, unit operation ontology, equipment ontology, material ontology and HAZOP ontology. Within each of the ontologies, concepts are organized in a hierarchy where concept nodes are connected by is-a links. Synonyms are given for concepts whose synonyms are available in the CPI. Process ontology is built based on the classification of chemical plant types such as hydrocracking plant, FCCU plant, ethylene plant, ammonia synthesis plant and so on. Process unit ontology and unit operation ontology are basically from OntoCape. In equipment ontology, equipment is specified by design properties such as design temperature, pressure and some structural specifications. Since MSDS information of each material is indispensable in HAZOP analysis, material ontology not only depicts the chemical species information, but also contains MSDS information for each material. In HAZOP ontology, HAZOP related conceptions such as nodes, parameters, guidewords, deviations, causes, consequences, safeguards, recommendations and risk are described. After the ontologies were created, the ontology editor Protégé (Stanford Medical Informatics, 2006) was used for verification.

3.3. Integrated reasoning framework for HAZOP analysis

As stated above, the existing HAZOP expert systems could only address "routine" and generic process analysis. Due to the lack of the ability of machine learning, they could not "remember" the analysis, especially "non-routine" analysis that has been done so that the HAZOP team has to do the analysis once again even if similar HAZOP analysis scenarios have been discussed before. To overcome this problem, a HAZOP expert system PetroHAZOP with learning capability is built by integrating CBR and the above six ontologies.

The HAZOP analysis process of the case-based expert system approach is illustrated in Fig. 3. PetroHAZOP consists of four modules, as shown in Fig. 4. In what follows, the four modules will be explained.

3.3.1. Case base module

Construction of the case base to a large extent determines the intelligence level of a CBR system. Each case instance generally consists of two parts: the problem and the solution. Inside the case base of PetroHAZOP, the problem part contains the HAZOP analysis background information of a particular deviation while the solution part describes its abnormal causes, adverse consequences, risk, safeguards, recommendations and some other auxiliary information such as the name of HAZOP team members and HAZOP analysis date. Stored in a relational database, each case holds a unique identification number. It is not uncommon that there are hundreds or

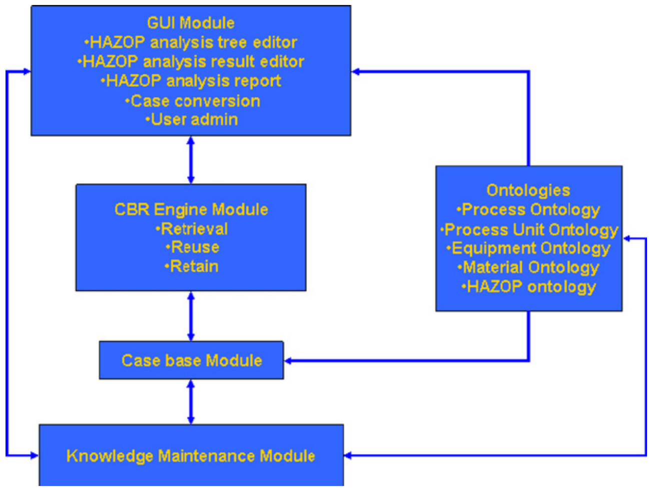


Fig. 4. Configuration of PetroHAZOP.

even thousands of deviations that need to be combed through during a HAZOP analysis of a typical chemical process. Therefore, the case base will grow to a very large size with time. To facilitate the similarity-based case retrieval that is described in the following section, a hierarchical case structure is introduced in this paper as an indexing method to partition a huge number of cases into multiple hierarchical subordinate case bases (SCB). HAZOP analysis cases can be categorized into the first level SCBs by the types of the chemical processes specified in the process ontology while those cases within a same type of chemical process can be further classified into the second level SCBs by the equipment types specified in the equipment ontology. Cases inside a second level SCB can still be divided into the third level SCBs according to the deviation types specified in the HAZOP ontology. This hierarchical case structure can be regarded as a knowledge-based indexing method where HAZOP domain-specific knowledge is applied, important features for quick and accurate retrieval of past cases (Barletta, 1991).

Each case in the case base is defined by indexes of four major categories: equipment with its design parameters, materials contained in the equipment, operating conditions, and stream context conditions. The equipment design parameters such as design pressure and design temperature describe the equipment where the deviation being analyzed occurs. The equipment type must be available in the equipment ontology. Each case contains a list of materials present in the equipment. Hazardousness related physical–chemical characters of materials such as flash point, boiling point and toxicity are distilled from MSDS to represent the material characters. Operating conditions include parameters such as operating temperature, pressure and level. The stream context conditions reflect the equipment types of both up stream and down stream of the equipment.

According to the standard IEC61882, the case solution which is the HAZOP analysis results should include information such as the deviation's root causes, adverse consequences, the safeguards available in the P&IDs, the recommendations for hazard mitigation.

3.3.2. CBR engine module

The CBR engine module (CEM) is the core of the indispensable phase of CBR systems, i.e. retrieval. When a new deviation analysis problem is presented to the system, the CBR engine is activated. The engine starts from selecting the corresponding SCB that fits the problem through the hierarchical indexing mechanism. Within the chosen SCB, all past cases are compared with the new problem, and scored based on the similarity-based case retrieval algorithm that is described in the following to find the closest-matching cases. To define the similarity between the past case and new problem, a measure is needed first to assess the closeness between the attributes belonging to them.

Basically there are five types of attributes for each case: object such as equipment, string such as the material name, numeric such as operating temperature of equipment, interval-numeric such as design parameters and set object such as materials. Fig. 5 shows different similarity algorithms that are employed in this paper to calculate the similarities of different types of attributes.

(1) The similarity of string attributes is simply calculated by string matching algorithm. If string attributes are same, then their similarity is 1, otherwise 0.
(2) In a HAZOP case, all numeric attributes are transformed to non-negative values. For example, the temperature unit is Kelvin temperature while the absolute pressure is used to represent pressure attributes.

    The mathematics similarity formula for non-negative numeric attributes x_i and y_i is:

$$\mathrm{sim}(x_i, y_i) = \frac{1 - |x_i - y_i|}{\max(x_i, y_i)} \qquad (1)$$

(3) Similarity between sets

    Usually, there is more than one material involved in a piece of process equipment. Comparison of HAZOP cases usually requires comparison of the material sets present in the equipments where the cases originate. We proposed a new approach

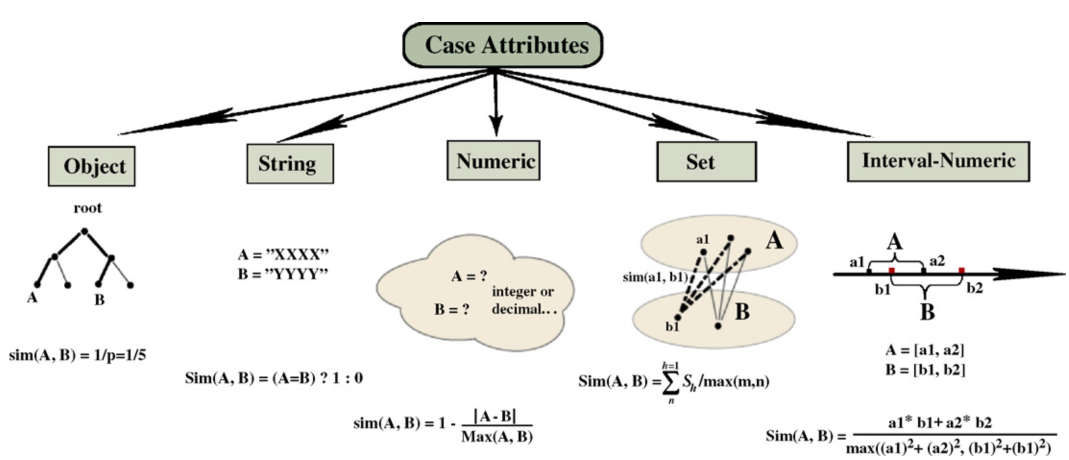


Fig. 5. Similarity algorithms for different types of case attributes.

(Set-Similarity method) to compute the material set similarity between different cases. The approach is expressed as follow.

Assume there are two material sets A and B, which belong to two different cases, respectively. Set A contains m materials: $M_{A1}, M_{A2}, \ldots, M_{Am}$, and set B contains n materials: $M_{B1}, M_{B2}, \ldots, M_{Bn}$. Then the similarity of A and B could be computed by Eq. (2).

$$\mathrm{sim}(A, B) = \frac{\sum_{i=1}^{N} S_i}{\max(m, n)} \tag{2}$$

where

$N = \mathrm{Min}(m, n)$

S_i is the maximum similarity between the ith material in one set and each material in the other material set, $1 \le i \le N$.

if $N = m$ then,

$$S_i = \underset{j=1}{\overset{n}{\mathrm{Max}}}\{sim(M_{Ai}, M_{Bj})\} \tag{3}$$

if $N = n$ then

$$S_i = \underset{j=1}{\overset{m}{\mathrm{Max}}}\{sim(M_{Aj}, M_{Bi})\} \tag{4}$$

In Eqs. (3) and (4), $\mathrm{sim}(M_{Ai}, M_{Bj})$ represents the similarity between material M_{Ai} and material M_{Bj}, which can be computed by Eq. (5)

$$\mathrm{sim}(M_{Ai}, M_{Bj}) = \sum_{k=1}^{K}(W_k \mathrm{sim}(att_{Aik}, att_{Bjk})) \tag{5}$$

where K represents the number of index attributes of a material, W_k represents the weight of the kth index attribute, $1 \le k \le K$, att_{Aik}, att_{Bik} are respectively the kth numeric index attributes of materials M_{Ai} and M_{Bj}, and $\mathrm{sim}(att_{Aik}, att_{Bik})$ can be calculated by Eq. (1).

(4) The similarity algorithm of interval-numeric feature is extended-Euclidian algorithm. Suppose there are two interval-numeric attributes $A = [a1, a2]$, $B = [b1, b2]$, then their similarity can be calculated as follows:

$$\mathrm{sim}(A, B) = \frac{(a_1 b_1 + a_2 b_2)}{\max((a1)^2 + (a2)^2, (b1)^2 + (b2)^2)} \tag{6}$$

(5) Object similarity

HAZOP cases have object attributes such as equipment and materials. The object similarity calculation takes advantage of the ontologies described in Section 3.2. Ontology based similarity algorithms have been reported in literature. In this paper, the path length measure is used to calculate the object similarity (Pedersen, Pakhomov, Patwardhan, & Chute, 2007). It essentially computes the similarity between two object nodes by counting the numbers of nodes on the shortest path between them in the ontology hierarchy. The shortest path includes both the object nodes. Mathematically, the similarity of two object nodes A and B using the path-length measure (path) is defined as:

$$\mathrm{sim}(A, B) = \frac{1}{p} \tag{7}$$

where p is the number of nodes on the shortest path between A and B within an ontology hierarchy.

For example, in the equipment ontology, if equipment A is a subclass of equipment P (subclass is equivalent to a is-a relationship in ontology), and equipment B is a subclass of equipment Q while equipment P and equipment Q are two subclasses of equipment O. The shortest path from equipment A to equipment B is A-P-O-Q-B. There are five nodes on the path. Therefore, the similarity of equipment A and equipment B is 1/5 (see Fig. 5).

Finally the case similarity is the sum of each case attribute similarity multiplied by its weight which is determined by domain experts. The weights are adjustable through the knowledge maintenance module which is described below.

3.3.3. Knowledge maintenance module

The effectiveness of a CBR system depends largely on the qualitative and quantitative richness of its stored cases which are the knowledge repository of past experiences. That is to say, the more quality cases stored in the case base, the more effectively the system reasons. Initially, there are few cases in the case base. Through the knowledge maintenance module (KMM), cases from the past HAZOP analysis records can be manually input to the subordinate case bases to facilitate CBR.

Case retaining is another feature of KMM. Revised cases can be converted to new cases and retained in the corresponding subordinate case bases. Before a new case is stored, term translation based on ontology is to be done if necessary. For example, "temperature too high" and "high temperature" will be translated into "more temperature" in the case base.

Since weight factors used in the similarity based case retrieval are predefined according to the knowledge of authors and a few expert consultants, it will not be surprised that there is a need to justify them according to the feedbacks of users in industrial practices. A user interface is designed in KMM to modify the weight factors through a certain level of authorization.

3.3.4. Graphical user interface (GUI) module

The GUI module contains graphical user interfaces to perform the following functions:

- Creating HAZOP analysis project
- Specifying equipment
- Specifying materials
- Selecting parameters and HAZOP guidewords
- Editing HAZOP analysis results
- Retrieving and reusing similar cases
- Retaining cases
- Reporting HAZOP analysis results
- Administrating users

Fig. 6 is a snapshot of a main GUI of PetroHAZOP. On its left-hand side is a treeview like the Microsoft's Windows Explorer tree. It displays a hierarchical collection of labeled items such as nodes, equipments, process variables and deviations. On its right-hand side is the area where causes, consequences, risk, safeguards, suggestions and comments for a certain deviation can be edited by users or automatically loaded from selected similar cases. Case Searching button on the top-right corner of the snapshot implies that the system is searching similar cases based on the specifications of the deviation being analyzed. Once the searching is done, the caption of button is changed to "Show similar cases". PetroHAZOP also has a user menu for project management, public data entry and user administration. On the bottom of the treeview, there are two more nodes, i.e. materials and report. In the materials node, the process materials can be specified by users. Customizable HAZOP reports can be generated by using the Report node.

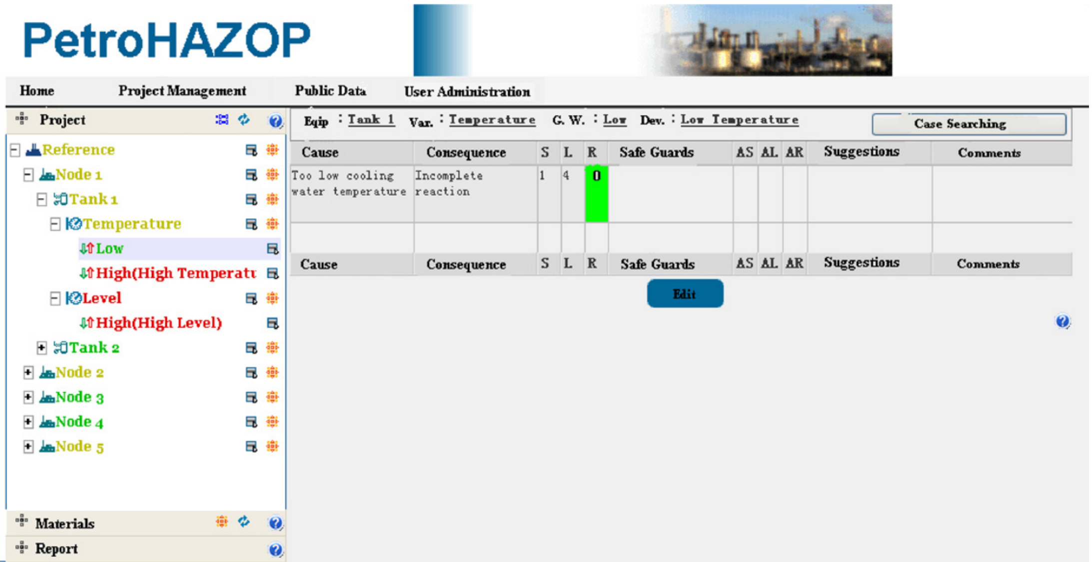


Fig. 6. Snapshot of a main graphical user interface of PetroHAZOP.

4. Application examples

PetroHAZOP is programmed with Java allowing concurrent multiple users to manipulate the system through intranet. This multiple user mode greatly improves the work efficiency of the HAZOP team. Recently, PetroHAZOP has been successfully installed and implemented at one of the largest oil companies in China. There were more than 900 cases in the case base at the time when this paper was written. The following examples demonstrate how the proposed HAZOP expert system PetroHAZOP can help with both "routine" and "non-routine" analysis.

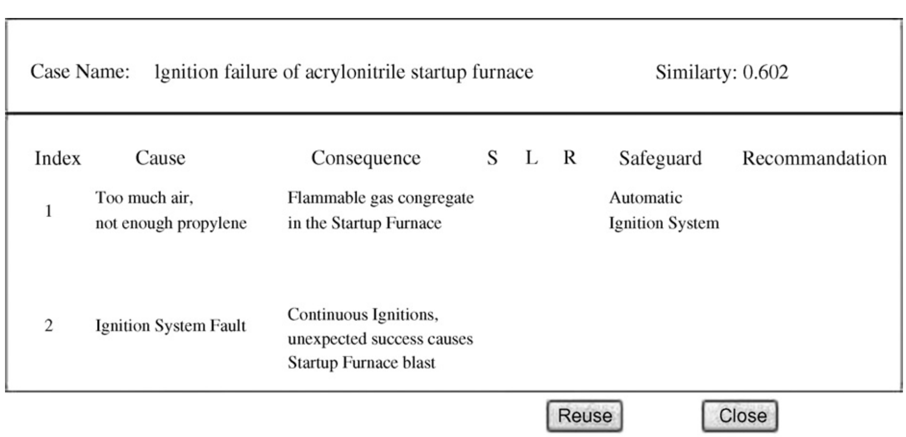
4.1. Example 1: "non-routine" analysis

Acrylonitrile production process is a highly hazardous process in the petrochemical industry. Most of the 16 materials involved in the process including raw materials, intermediate products, products and by-products are flammable, toxic or/and volatile. There are about 14 major equipments such as reactor, chilling tower, absorption column, recycle column, distillation columns. The whole process was divided into six nodes. Totally 87 deviations of 59 key parameters had been analyzed by a HAZOP team. We transfer the results through the knowledge maintenance module into cases, resulting in 87 cases in the case base.

The vinyl chloride production process (VCPP) is another highly hazardous chemical process which consists of three sections: chlorine/hydrogen processing section (CHPS), hydrochloride synthesis section (HSS) and vinyl chloride synthesis section (VCSS). One example of the 'non-routine' analysis of this process was 'ignition failure of the HCl synthesis furnace' since it is hard to be modeled or generalized with rules. With PetroHAZOP, a similar case 'ignition failure of acrylonitrile startup furnace' from the case base was automatically retrieved with the similarity degree of 0.602. The found case is analogous in the equipment type in the equipment ontology and the deviation type in the HAZOP ontology. If the user clicks on the button "Show similar cases", he/she can find out the causes and consequences of similar case (Fig. 7). This similar case can be reused if the button "Reuse" on Fig. 7 is pressed. Here reuse means that the causes and consequences are automatically loaded to the causes and consequences of the new case. The user then can edit the causes and consequences if necessary.

4.2. Example 2: "routine" analysis

Another example is the completeness checking even for routine HAZOP analysis by means of CBR. A HAZOP team was assigned to HAZOP a polymer production process. Node 1 containing a stirred tank was analyzed first and it took the team about 2 days to


Case Name: Ignition failure of acrylonitrile startup furnace Similarty: 0.602

Index	Cause	Consequence	S	L	R	Safeguard	Recommandation
1	Too much air, not enough propylene	Flammable gas congregate in the Startup Furnace				Automatic Ignition System	
2	Ignition System Fault	Continuous Ignitions, unexpected success causes Startup Furnace blast					

Fig. 7. Snapshot of case 'Ignition failure of acrylonitrile startup furnace'.

complete its analysis. In the third day, the team started the analysis of Node 2 which also contained a stirred tank. When the team was about to close the analysis of the 'Low Pressure' deviation in the stirred tank of Node 2, the leader asked the recorder who was responsible for manipulating PetroHAZOP to retrieve similar cases to check if anything was missed. A similar case 'Low pressure in the stirred tank of Node 1' was discovered in the case base and one of its causes, 'axial sealing leak', was not considered for the current case. The consequence of the cause was "oxygen entering the stirred tank leading to contamination and deactivity of the catalyst in the tank". Since the cause and consequence could also happen in the current case, they decided to reuse the found case, and minor modification such as change of the catalyst name was made for the new case.

5. Conclusions

HAZOP analysis requires high accuracy, consistency and completeness because any ignorance would lead to catastrophic losses. Therefore, the HAZOP team must ensure that it would not lose any resources that are available to help them meet the above requirements. As a solution, this article offers an integrated solution for the complex problems in the path of automating HAZOP. The proposed HAZOP expert system PetroHAZOP not only facilitate "routine" analysis but also "non-routine" analysis due to its learning capability by which the HAZOP analysis quality can be continuously improved during practice. Due to the adoption of CBR mechanism, PetroHAZOP can map past experiences to the new cases. Therefore, it is more adaptive in HAZOP analysis and greatly eases the knowledge management and knowledge dissemination. Even though HAZOP has been widely practiced in the CPI of developed countries, it just starts to get recognized by the practitioners in China. Lack of HAZOP experts hinders the wide implementation of HAZOP in the chemical plants. Hopefully, PetroHAZOP can facilitate the industrial exercises of HAZOP in China and contribute to the loss prevention in the largest developing country where chemical accidents are posing a serious threat to its fast development.

Automatic adaptation for HAZOP analysis is an outstanding task that could be addressed by introducing other artificial intelligence technologies to CBR. Layered digraph model (LDG) has been proposed by authors to perform both "non-routine" and routine HAZOP analysis (Cui, Zhao, Qiu, & Chen, 2008). Future work will be oriented to integrate CRB with LDG based reasoning.

Acknowledgements

This work is supported by Program for New Century Excellent Talents in University. Anonymous reviewers of this paper are highly appreciated for their helpful comments and suggestions.

References

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues. Methodological variations, and system approaches. *Artificial Intelligence Communications*, *7*(1), 39–59.

Aha, D. W. (1998). The omnipresence of case-based reasoning in science and application. *Knowledge-Based Systems*, *11*(5–6), 261–273.

Barletta, B. (1991). An introduction to case-based reasoning. *AI Expert*, *6*(8), 42–49.

Batres, R., West, M., Leal, D., Price, D., Masaki, K., Shimada, Y., et al. (2007). An upper ontology based on ISO 15926. *Computers & Chemical Engineering*, *31*(5–6), 519–534.

Cui, L., Zhao, J., Qiu, T., & Chen, B. (2008). Layered digraph model for HAZOP analysis of chemical processes. *Process Safety Progress*, *27*(4), 293–305.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–221.

Kolodner, J. (1993). *Case-based reasoning [M]*. San mateo, CA: Morgan Kaufmann Publishers, Inc.

López-Arévalo, I., Bañares-Alcántara, R., Aldea, A., Rodríguez-Martínez, A., & Jiménez, L. (2007). Generation of process alternatives using abstract models and case-based reasoning. *Computers & Chemical Engineering*, *31*(8), 902–918.

McCoy, S. A., Wakeman, S. J., Larkin, F. D., Jefferson, M. L., Chung, P. W. H., Rushton, A. G., et al. (1999). HAZID, a computer aid for hazard identification 1. The STOPHAZ package and the HAZID code: An overview, the issues and the structure. *Process Safety and Environmental Protection*, *77*(6), 317–327.

McCoy, S. A., Wakeman, S. J., Larkin, M. L., Chung, P. W. H., & Rushton, A. G. (2000). *HAZID, a computer aid for hazard identification: 4. Learning set, main study system, output quality and validation trials. Process Safety and Environmental Protection*, *78*(2), 91–119.

Morbach, J., Yang, A., & Marquardt, W. (2007). OntoCape—A large-scale ontology for chemical process engineering. *Engineering Applications of Artificial Intelligence*, *20*(2), 147–161.

Pedersen, T., Pakhomov, S. V. S., Patwardhan, S., & Chute, C. (2007). Measure of semantic similarity and relatedness in biomedical domain. *Journal of Biomedical Informatics*, *40*(3), 288–299.

Stanford Medical Informatics. (2006). *The protégé ontology editor and knowledge acquisition system*. Available at http://protege.stanford.edu.

Venkatasubramanian, V., Zhao, J., & Viswanathan, S. (2000). Intelligent systems for HAZOP analysis of complex process plants. *Computers & Chemical Engineering*, *24*(9–10), 2291–2302.

Zhao, C., Bhushan, M., & Venkatasubramanian, V. (2003). Roles of ontology in automated process safety analysis. *Computer Aided Chemical Engineering*, *14*, 341–346.

Zhao, C., Bhushan, M., & Venkatasubramanian, V. (2005). PHASUITE: An automated HAZOP analysis tool for chemical processes Part I: Knowledge Engineering Framework. *Process Safety and Environmental Protection*, *83*(B6), 509–532.