

Proposition of ontology for safety using cased based reasoning in industrial context

Nassima AISSANI, Hadj Mohamed Islam GUETARNI

Abstract— Safety is seen as a key factor for a successful business and an inherent element of business performance. Safety control means changing system design parameter as well as production parameters and human behaviour to meet safety requirements. It is also necessary to face the risks and accidents that can occur. In this paper, the aim is to present ontology for risk assessment in industrial case to feed a case based reasoning data base in order to evaluate risks and then regulate human behavior and the system parameters;

Keywords— Safety control; Ontology; Sementic Network; case based reasoning

I. Introduction

The increasing complexity and size of electronic systems in industry, joint with the growing market demand, requires from the industry to implement an efficient safety system to preserve equipment viability, environment and especially human life protection. Often physical and virtual barriers are put in place but it remains insufficient. In 1985 Wallace and De Balogh [1] pointed the importance of using DSS (Decisional Support Systems) for disaster management, by taking into consideration a DSS as well as an organized database which can be asked in different manners and which can give a suitable answer establishing a safety policy in toxic spill management. Considering a real industry case, the aim is to develop a DSS able to predict risk and propose actions to execute in critical situation. This system is based on case based reasoning approach and the aim in this paper is to present an ontology which will be used to feed the case base and to retrieve cases. In the next section, a simplified review of DSS use in safety area will be presented.

II. DSS in risk evaluation and crisis recovery

A. Risk evaluation

DSS are, already, used to determine dynamically the risk of a considered situation, for example evaluating risks caused on roads by the transportation of hazardous materials, by taking into account weather, driver, state of the road, etc [2].

DSS are also used It is also used in environment prevention, [3] used it to evaluate the risk of pollution due to the use of pesticides considering: Groundwater risk value, Air risk value, Soil risk value, Surface water risk value including environmental conditions, pesticide properties, and pesticide application records. The developed DSS (PURE) is available on web for guiding growers to select reduced risk pesticides to use in their crops (at <http://pure.ucdavis.edu>).

B. Reponses in crisis

DSSs are also used to suggest actions to be fulfilled in crisis situations, dealing with emergency, evaluating risk and proposing alternatives. Most of these applications are based on an important knowledge base. The main problem in this kind of application is the duration of knowledge base consultation and the development of response which are often important. For example, [4] used augmented reality to develop this knowledge base to develop DSS in Housing Health area. [5] Used a network model for data acquisition that feed the DSS, this network is composed of information nodes: experts, geografical system.

In this paper, ontology is presented to structure the data and retrieve case for safety system based on case based reasoning appraoch used for risk evaluation. In the next section, the case of study will be presented to developp the corresponding ontology.

III. DSS architecture for risk evaluation

The architecture components of DSS consist of knowledge base, inference/reasoning engine, and user communication/interaction. Figure 1 shows the architecture components of DSS. Whereas the knowledge base is made up guidelines, rules, and probabilistic models, the inference/reasoning engine combines the data in the knowledge base with that of the new case description. The user communication component of the architecture consists of a simple way of getting data into the system and getting results/ risk evaluation to the user. The fact that the architecture of the DSS depends on knowledge bases means that inappropriate representation of data, information, and knowledge present enormous threats to the adoption of DSS in risk analysis practice. It's why, in this paper the aim is to present a rigorous manner to present and store cases and then retrieve that using ontology.

Nassima AISSANI, Hadj Mohamed Islam GUETARNI

Laboratoire de l'Ingénierie de la sécurité industrielle et du développement durable
Institut de Maintenance et de Sécurité Industrielle
University of Oran 2

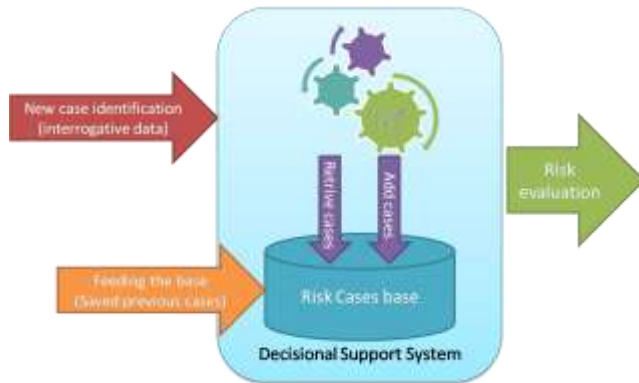


Figure. 1. Decisional support system architecture

iv. Case of study and risk analysis

Our case study is a Cutting Industrial machine in textile industry. This machine is the most critical engine in the factory, used to cut the canvas. The breakdown of this machine causes the degradation of the whole system reliability, i.e. the entire factory by 80%. In the present study, a focus is dedicated to this machine. Developing an appropriate case representation is one of the most critical issues in a CBR (case based reasoning) system. Representation should identify the unique characteristics of a case in both problem description and solution parts. Risk events of cutting machine should be recorded as well as their precursor. Initially, risk analysis should be done.

A. Event tree elaboration

Structured, rigorous, methodical approach and permits probability assessment for risk evaluation. Many event trees are developed for many dreaded events but one of them is presented here as example. “The worker's injury” Is one of the most feared events. The event tree is analyzed to get maximum information on a dreaded event .As seen on figure 2, causes are especially of two types: Human fault and dysfunction of the system equipment.

With 30 events and 9 basic events, 7 minimal cuts set are developed. Over 100000 operating hours: The probability of the dreaded event is $1,65 \cdot 10^{-5}$. There are 10 minimal cut sets from 1 to 3 levels. The used probabilities distributions are: Constant, Exponential distribution, Weibull distribution and General Linear Model. Using this information a semantic network and then ontology could be developed..

B. Ontology developpmt

Risk evaluation and prevention require and rely on extensive knowledge about the production system architecture, hardware and software components, features, characteristics and specificities, normal and abnormal operating conditions, interdependencies, performance, etc..

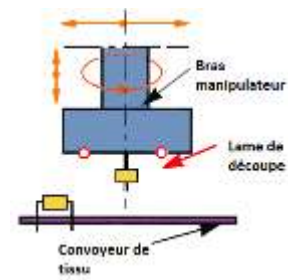


Figure. 3. Cut machine and its parameters

Based on the investigation and analysis of this cutting machine (figure 3), the incident categories are defined, as Motor Failure, DC motor in our case (MF), Dysfunction of Accessories (DF) and Human Fault (HF) relevant to improper handling or repair error, for example: Maintenance performed poorly or not. The trees of events have made it possible to link these precursors to the risks with a hierarchization (Fig 4).

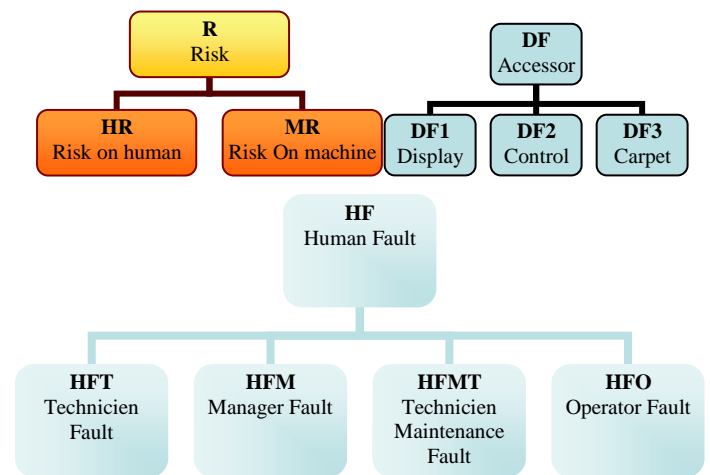


Figure. 4. Hierarchy of faults and precursors

Using a semantic relationship to connect the “Precursors” concept and “Safety risk” concept to the “Safety measures” concept can ensure the potential risk. Recommended safety measures will not be ignored once the “Precursors” concept and “Safety risk” concept are identified. The hierarchization will help to develop the corresponding ontology (Fig 5). Protégé was used as construction tool to assist the design (<http://protege.stanford.edu/>)

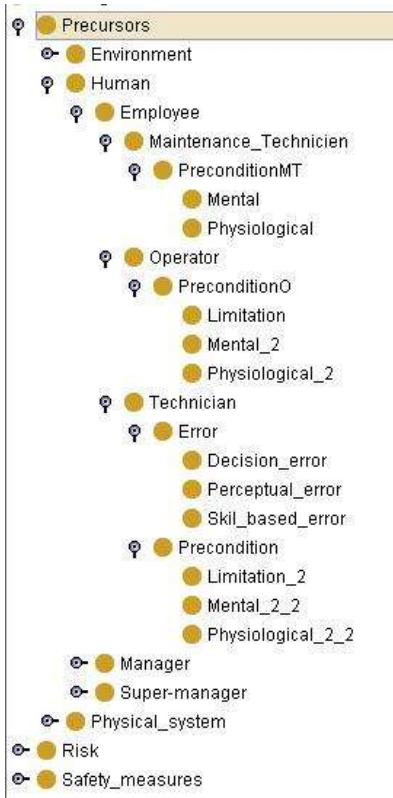


Figure 5. Precursors, risk and safety mesures

The risk “Injury of the Worker” is an instance of the human-risk class as shown in (Fig 6) the deduced probability from the event tree is its value.

Figure 7 shows the structure of the risk-class in terms of relations with derivable classes namely: Human-risk and Machine-risk. In the next section case retrieval algorithm based on ontology will be presented.



Figure 7. structure of classes

v. Retrieval of cases and similarities

The quality of retrieval approaches directly affects the relevance of retrieved cases in order to identify the cases in the database approaching the case posed. In case retrieval, similarity scoring (or distance-based) is commonly used, such as taxonomy tree approach and use of tables or rules [6]. Another possible approach is to adopt a mix of approaches for different types of attributes; [7] used five different similarity scoring functions. In this research, the similarity scoring approach makes use of a semantic network to calculate the

similarity score between two precursors. The basic idea of the proposed similarity scoring approach is that the higher the proportion of common sub-concepts two values have, the more similar the two values are. Similarity calculation between cases has two steps. The first step is to calculate local similarity between precursors where a weight of sub-concepts is assigned to each of the nodes. The guiding principle to assign weights is that higher nodes which are closer to the root are more influential on the categorization of precursor (fig 5). Thus the similarity due to a match on the sub-concept can be calculated with these weights. Suppose that case C_1 has precursors $pr_1, pr_2, \dots, pr_i, pr_{n1}$ and C_2 has precursors $PR_1, PR_2, \dots, PR_j, PR_{m1}$ with variable “Human”; case C_1 has precursors $q_1, q_2, \dots, q_i, q_{n2}$ and C_2 has precursors $Qs_1, Qs_2, \dots, Qs_j, Qs_{m2}$ with situation variable “Physical-system”; case C_1 has precursors $rs_1, rs_2, \dots, rs_i, rs_{n3}$ and C_2 has precursors $Rs_1, Rs_2, \dots, Rs_j, Rs_{m3}$ with variable “environment”.

The local similarity between precursor’s pr_x in case C_1 and PR_z in case C_2 with the same situation variable can be calculated based on the following equation:

$$Simil(pr_x, PR_x) = \frac{\sum_{i=1} \theta_{ci}}{\left(\sum_{i=1}^a \theta_{ci} + \sum_{j=1}^b \theta_{dj} \right)}$$

(0.1)

Where $i=1..a, j=1..b$ are the number of concepts that belongs only to either pr_x or PR_z ; θ_{ci} = weight of the common concept i , and θ_{dj} = weight of the concept j that belongs only to either pr_x or PR_z . $Simil(pr_x, PR_z)$ measures the proportion of weights represented by the common concepts to the weights of both concepts . The value of proportion 1 means that concepts are exactly identical. Then, the global similarity is computed based on a weighed sum of the local similarity of all precursors’ pairs of cases C_1 and C_2 being compared, as follows:

$$simil(C_1, C_2) = \sum_{z=1}^{m1} \max \{ simil(pr_1, PR_z), \dots, simil(pr_{n1}, PR_z) \} w_{pri} + \sum_{z=1}^{m2} \max \{ simil(q_1, Qs_z), \dots, simil(q_{n2}, Qs_z) \} w_{qi} + \sum_{z=1}^{m3} \max \{ simil(rs_1, Rs_z), \dots, simil(rs_{n3}, Rs_z) \} w_{rsi}$$

(0.2)

Where: $\max \{ simil(pr_1, PR_z), \dots, simil(pr_{n1}, PR_z) \}$

is the effective value of the local similarity for the precursor PR_z with the case C_1 , w_{pri} = corresponding weight of the precursor pr_i which is the most similar with the precursor PR_z . The value of w_{pri} reflects the relative importance of the precursor pr_i to the accident of the case C_1 . Because of the limited quantity of cases, the frequency of the precursor pr_i appearing in the same risk events cannot prove the importance of the precursor pr_i . So the corresponding weight of the

precursor can be calculated as an average value based on the following equation:

$$w_{pri} = w_{qst} = w_{rst} = \frac{1}{\max(m1, n1) + \max(m2, n2) + \max(m3, n3)} \quad (0.3)$$

VI. Experimentation

The developed ontology has given rise to a java model composed of class generator of each developed concept in this ontology (fig 8). This is facilitated by using the Jena library (<https://semtools.ecoinformatics.org/repository/dev/sms/lib/jena.jar/view>). Then by using the ontology metrics and according each concept position in ontology a score is given to each class to be used in case retrieval.

But the similarity measures are not always applicable to each pair of concepts (or instances) of the ontology. For example, environment precursor must not be compared at all with human safety measures. To prevent such comparisons, notion of similarity regions is defined. A similarity region is a sub-branch of the ontology hierarchy where concepts and instances are comparable with each other (Fig 9). The definition of such regions is manual and depends on the target application, so a variable is added to identify each region. To compute the similarity between a query attribute and a case attribute, it must be verified first if these two attributes belong to the same similarity region. This developed ontology allow description and save about 172 “worker injury” cases and 1806 other risks cases in the DSS case base. The next step is to develop the inference engine and test the retrieval case method to validate its efficiency.

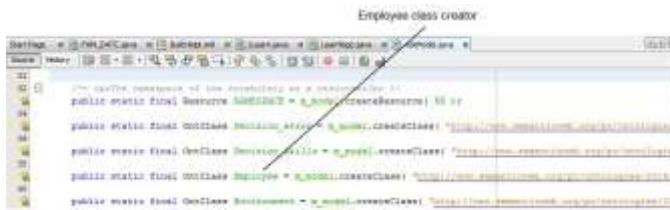


Figure 8. Ontology class creation

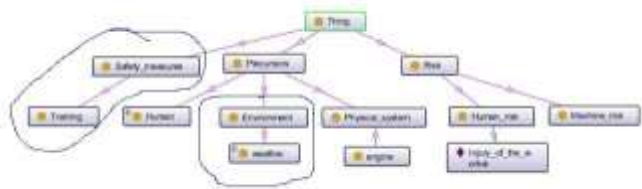


Figure 9. Similarity regions

VII. Conclusion

The presented paper is a proposition of safety ontology for an industrial case study. This ontology is primordial to capitalize and organize knowledge about the risks associated to this case, their precursors, the security measures and evaluation. This ontology will be used to identify the unique characteristics of a case in both problem description and solution/evaluation parts in DSS (Decisional support system) based on CBR (case based reasoning). Case retrieval algorithm based on scores dedicated from ontology is also proposed. In experimentation phase, similarity regions notion was also introduced to be pickier in comparing cases. As a perspective, the developed case base should be combined to the inference engine in order to test its effectiveness and the efficiency of the retrieval case method, the approach should also be compared with others used retrieval case algorithms known from the literature.

References

- [1] Wallace, W. A., Balogh, F. D., 1985. Decision Support Systems for Disaster Management. Public Administration Review, 45, pp. 134-146
- [2] Giglio, D., Minciardi, R., Pizzorni, D., Rudari, R., Sacile, R., Tomasoni, A., and Trasforini, E. (2004). Towards a decision support system for real time risk assessment of hazardous material transport on road. Proceeding IEMSS 1-6
- [3] Zhan, Yu, et Minghua Zhang. 2012. « PURE: A Web-Based Decision Support System to Evaluate Pesticide Environmental Risk for Sustainable Pest Management Practices in California ». Ecotoxicology and Environmental Safety 82 (août): 104-113. doi:10.1016/j.ecoenv.2012.05.019.
- [4] Daouda Kamissoko, Pascale Zaraté, François Pérés. Decision process in large-scale crisis management. Environment Systems & Decisions, 2014, vol. 34 (n 2), pp. 277-287
- [5] Kaklauskas, A., Zavadskas, E. K., Cerkauskas, J., Ubarte, I., Banaitis, A., Krutinis, M., & Naimaviciene, J. (2015). Housing Health and Safety Decision Support System with Augmented Reality. Procedia Engineering, 122, 143-150.
- [6] Huysmans, J, Karel Dejaeger, Christophe Mues, Jan Vanthienen, and Bart Baesens. An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models. Decision Support Systems , 51(1):141-154, 2011.
- [7] Zhao F., et al. A probabilistic and continuous model of protein conformational space for template-free modeling. J. Comput. Biol. 2010a;17:783-798

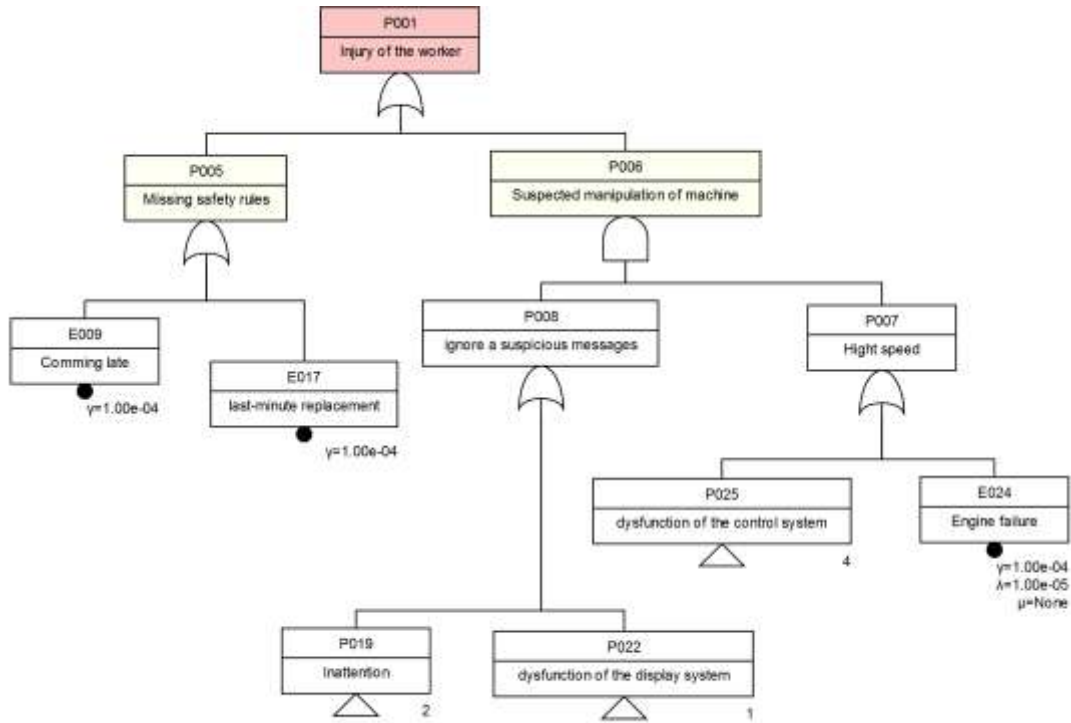


Figure 2. Event tree "Injury of the worker"

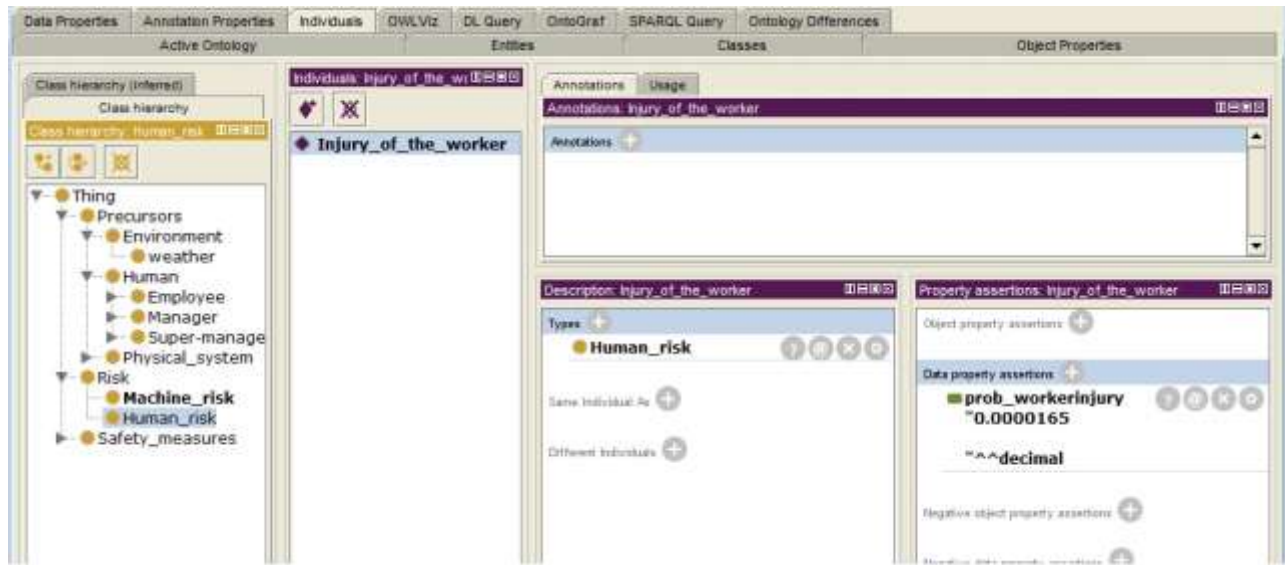


Figure 6. Instanciation of Human risk class by the risk "Injury of the worker"