Case-Based Reasoning Approach to Construction Safety Hazard Identification: Adaptation and Utilization

Y. M. Goh[1] and D. K. H. Chua, M.ASCE[2]

Abstract: Risk assessment, consisting of hazard identification and risk analysis, is an important process that can prevent costly incidents. However, due to operational pressures and lack of construction experience, risk assessments are frequently poorly conducted. In order to improve the quality of risk assessments in the construction industry, it is important to explore the use of artificial intelligence methods to ensure that the process is efficient and at the same time thorough. This paper describes the adaptation process of a case-based reasoning (CBR) approach for construction safety hazard identification. The CBR approach aims to utilize past knowledge in the form of past hazard identification and incident cases to improve the efficiency and quality of new hazard identification. The overall approach and retrieval mechanism are described in earlier papers. This paper is focused on the adaptation process for hazard identification. Using the proposed CBR approach, for a new work scenario (the input case), a most relevant hazard identification tree and a set of incident cases will be retrieved to facilitate hazard identification. However, not all information contained in these cases are relevant. Thus, less relevant information has to be pruned off and all the retrieved information has to be integrated into a hazard identification tree. The proposed adaptation is conducted in three steps: (1) pruning of the retrieved hazard identification tree; (2) pruning of the incident cases; and (3) insertion of incident cases into the hazard identification tree. The adaptation process is based on the calculation of similarity scores of indexes. A case study based on actual hazard identifications and incident cases is used to validate the feasibility of the proposed adaptation techniques.

DOI: 10.1061/(ASCE)CO.1943-7862.0000116

CE Database subject headings: Construction management; Accidents; Hazards; Safety; Risk management; Information management; Knowledge-based systems.

Author keywords: Accidents; Hazard identification; Safety; Risk assessment; Safety management system; Knowledge management; Case-based reasoning.

Introduction

Hazard identification is a core process in safety planning (British Standards Institute 2007). However, the effectiveness of hazard identification can be easily influenced by the risk attitude of individuals and the group conducting the risk assessment (Hillson and Murray-Webster 2007). In addition, either due to lack of experience or lack of time, hazard identification frequently fails to identify hazards associated to an activity (Gadd et al. 2004). It is suggested that one of the ways to reduce these potential problems of hazard identification is to utilize artificial intelligence (AI) tools to facilitate the hazard identification process. Even though AI tools are currently not able and not intended to replace human judgment, it is able to improve the efficiency and comprehensiveness of the hazard identification task.

Despite its potential benefits, research on the application of AI tools on hazard identification and risk assessment is very lacking. One of the few relevant researches is the attempt of Balducelli and D'Esposito (2000) to develop an AI approach based on case-based reasoning (CBR), genetic algorithm and numerical simulation to assist in the risk management and planning of fire emergencies. Another study by Mendes et al. (2003) produced a CBR approach to offshore well design that takes into account the risk of the design. Even though these studies are not meant to improve the efficiency and quality of hazard identification per se, it can be observed that CBR and AI as a whole has the potential to achieve this.

To improve the standard of construction safety risk assessment, which consists of hazard identification and risk analysis (British Standards Institute 2007), Goh and Chua (2009) presented a new CBR approach that utilizes records of past incidents and hazard identifications to facilitate new hazard identification processes. The proposed CBR approach aims to facilitate feedback of past experiences (Chua and Goh 2004) so that hazard identification is more comprehensive and efficient. This paper follows the same work and presents the adaptation mechanism for the CBR approach. A threefold strategy is adopted to derive an adapted hazard identification tree that incorporates past experience that is relevant to the presented case. The approach is demonstrated and validated with a case study using an existing database of incident cases and hazard identification documents. As in a large number of CBR approaches, the adaptation strategy presented herein is unique and has the potential to improve real world processes.

[1]Senior Lecturer (Occupational Health and Safety), School of Public Health, Curtin Univ. of Technology, GPO Box U1987, Perth, Western Australia 6845, Australia (corresponding author). E-mail: Y.goh@curtin.edu.au

[2]Associate Professor, Dept. of Civil Engineering, National Univ. of Singapore, 119260, Singapore. E-mail: cvedavid@nus.edu.sg

Note. This manuscript was submitted on September 18, 2008; approved on July 2, 2009; published online on July 4, 2009. Discussion period open until July 1, 2010; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 136, No. 2, February 1, 2010. ©ASCE, ISSN 0733-9364/2010/2-170–178/$25.00.

J. Constr. Eng. Manage., 2010, 136(2): 170-178

Case-Based Reasoning Approach

CBR is based on the basic paradigm that humans solve new problems by recalling past experiences (Mount and Liao 2001). The three key processes of CBR are: (1) case representation and indexing; (2) retrieval of cases; and (3) case utilization and adaptation (Chua et al. 2001; Kolodner 1993). Case representation and indexing involves the codifying of lessons that a case teaches and the context in which the case can teach its lessons. The retrieval of cases is the process of searching and determining relevance of past cases in the case base. CBR utilizes an intelligent and fuzzy methodology so that past knowledge can be retrieved even if the new situation is not exactly the same as the current situation. Case utilization and adaptation refers to the process of making changes to the retrieved cases to suit the new situation and harnessing the retrieved knowledge to meet the purpose of the users.

The proposed CBR hazard identification approach is based on the job hazard analysis (also known as job safety analysis) (Swartz 2002) approach. To facilitate retrieval, adaptation, and utilization of cases, a knowledge representation scheme derived from the modified loss causation model (Chua and Goh 2004) was adopted. One of the key components of the knowledge representation scheme is the indexing vocabulary that describes the situational variables or key characteristics of a job step or incident. The indexing vocabulary has the following linguistic structure: *Action(s)* executed on *object(s)-worked-on* using *resource(s)* at *location(s)* with *nearby object(s)* and *nearby action(s)*.

Another key component of the knowledge representation scheme is the representation of the sequence of incident events or incident sequence. The incident sequence comprises the consequence caused by a contact event which resulted from a breakdown event. Incident cases contain one incident sequence, while hazard identification trees consist of a set of inverted incident sequences identified by the risk assessment team. A set of situational variables describes the context for the incident case or the hazard identification tree, and forms the root node.

Goh and Chua (2009) presented the retrieval mechanism which is based on a subconcept approach that uses a semantic network (Russell and Norvig 1995), instead of a taxonomy tree as used by Kolodner (1993). A semantic network with weighted subconcepts is constructed for each situational variable and local similarity score (LSS) of a situational variable is calculated based on the weighted fraction of common subconcepts over the sum of common and different subconcepts. The adaptation strategy, which will be elaborated later, is also based on LSS. GSS (Global similarity score) is the overall similarity score between two cases calculated based on weighted LSS. In the proposed approach, a hazard identification tree with the highest GSS and a set of incident cases with GSS higher than a user-specified threshold value (default value of 0.6) will be retrieved. The similarity approach, in contrast to the exact matching approach used in traditional databases, allows similar cases to be retrieved and ensure that important information on hazards and risks associated with these cases are presented to the new case. Users can adjust the threshold similarity score to achieve a balance between relevance and number of retrieved cases. When the threshold is increased the relevance can be increased, but the number of cases will be reduced. Since a similarity score of 0.5 represents “neither similar nor dissimilar,” the default value of 0.6 was deemed to be a reasonable threshold for a good compromise.

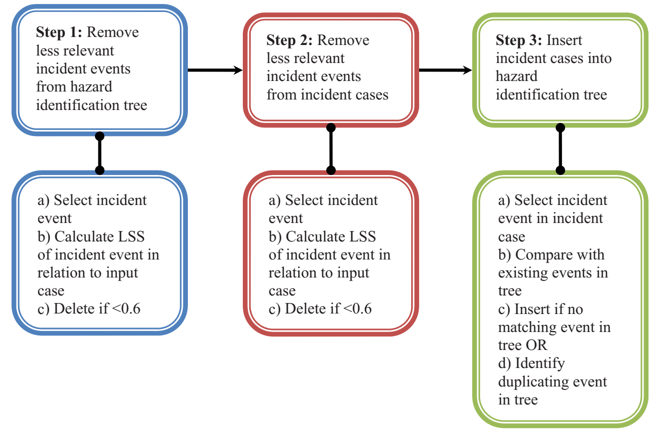

Fig. 1. Adaptation process

Adaptation Strategy

Even though the retrieved cases are deemed to be similar to the set of situational variables describing the job step being assessed, they may contain less relevant portions that should be adapted to the new work scenario (input case). Kolodner (1993) proposed ten adaptation strategies, which are classified under substitution, transformation, and other methods. It is possible to implement the various types of adaptation strategies using some form of rules (Bergmann and Wilke 1998). For instance, the CBR Works development shell, Empolis Knowledge Management GmbH (2001) makes use of rules with a set of preconditions and corresponding conclusions to execute the adaptations. Such rule-based adaptation is widely applied in different CBR systems (CBRS), for example Luu et al. (2003), Suh et al. (1998), and Sinha and May (1996).

Adaptation is a generic and high level process that can be implemented through any rational and sound method. Other strategies have been tried that differ from those proposed by Kolodner (1993). For instance numerous CBRS have used different mathematical or statistical formulae for adaptation purposes (Chua et al. 2001; Suh et al. 1998). Other tools like genetic algorithm, decision tree, neural network, and statistical models can also be implemented to adapt the retrieved cases.

To ensure proper adaptation, two key tasks must be accomplished. First, portions of the retrieved cases that are not applicable should be removed. Second, the retrieved incident cases will have to be incorporated into the hazard identification tree to produce a single adapted hazard identification tree that is relevant to the case in view. In this paper, the adaptation is based on the computations of the LSS at the nodes of the retrieved hazard identification tree or incident sequence to guide the deletion or insertion of incident events so as to derive an adapted hazard identification tree.

The proposed hazard identification adaptation belongs to the transformation type (Kolodner 1993), but unlike classical transformation methods, the transformation adaptation employed is executed in three sequential steps (see Fig. 1). The first step focuses on the retrieved hazard identification tree, where the hazard identification tree is pruned to delete irrelevant incident events. The second and third steps focus on the set of retrieved incident cases, where the retrieved incident cases are trimmed to remove irrelevant incident events (Step 2) and then integrated into the hazard identification tree (Step 3).

J. Constr. Eng. Manage., 2010, 136(2): 170-178

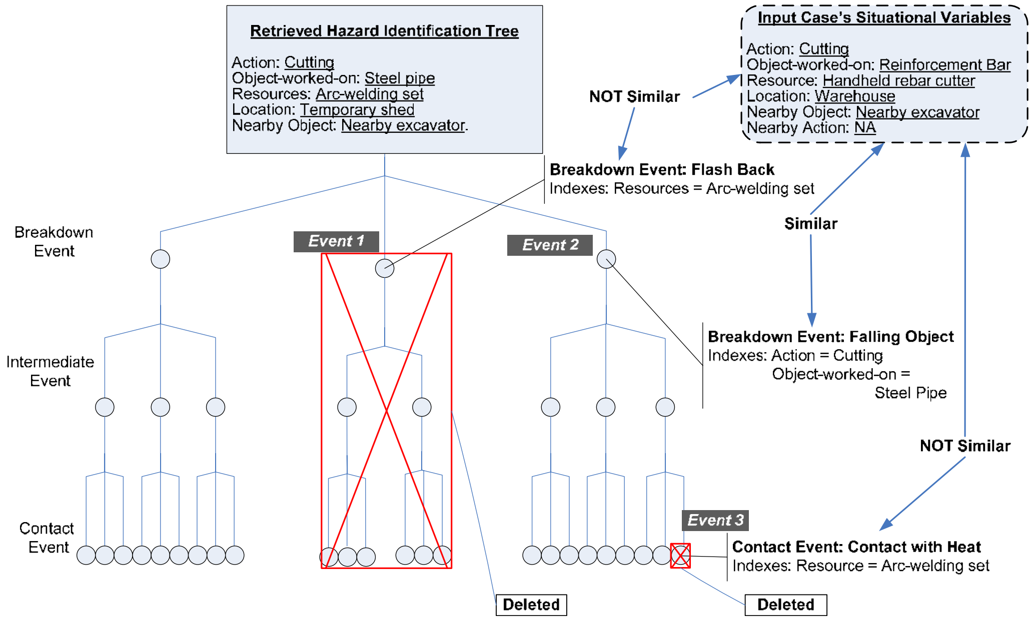

Fig. 2. Example of a hazard identification tree being pruned

Pruning of Retrieved Hazard Identification Tree

The first step in the adaptation process is to prune the retrieved hazard identification tree of any irrelevant incident nodes. This adaptation is based on the comparison of the indexes of individual incident events (see indexing vocabulary presented earlier) in the hazard identification tree with the set of situational variables of the input case. The indexes of incident events (breakdown, intermediate, or contact events) are user-selected subsets of the indexes of the retrieved case which are assigned during the original codification of the hazard identification tree. The assignment is based on the relevance of the index to the specific incident event. For example, in the retrieved hazard identification tree of Fig. 2, Event 2 is a Breakdown Event that has associated with it the indexes Action=Cutting and Object-worked-on=Steel Pipe, which also is a subset of the indexes of the retrieved hazard identification tree (shown at the root node). These indexes were deemed relevant to the event because the object that could fall off was the steel pipe while it was being cut without adequate support.

During adaptation, if any of the indexes of an incident event is *not similar* to the corresponding situational variable of the input case, the incident event will be deleted together with all the other events in its child nodes. As discussed earlier, *similarity* is measured using the LSS and a default threshold value of *0.6* (user specified) is used to identify similar cases. When the threshold similarity value is increased, more event nodes can be expected to be pruned. As discussed earlier, this threshold can be adjusted to balance the relevance of the incidents and the number of relevant event nodes to retain.

The LSS between index values V_1 and V_2 of the event node and the input case, respectively, is computed based on the similarity of the subconcepts representing the index values. Essentially, the LSS measures the proportion of weights represented by the common subconcepts to the weights of the union of the subconcepts, so that the higher the proportion, the greater will be the similarity of V_1 and V_2. Specifically, it is calculated as follows:

$$\mathrm{LSS}(V_1,V_2)=\sum w_{ci}\Big/\left(\sum w_{ci}+\sum w_{dj}\right) \qquad (1)$$

where $i=1,2,\ldots common$; $j=1,2,\ldots different$, *common* is the number of common subconcepts to both V_1 and V_2, *different* is the number of subconcepts that belongs only to either V_1 or V_2, w_{ci} the weight of the common subconcept i, and w_{dj} the weight of the subconcept j that belongs only to either V_1 or V_2. The details for the similarity scoring and the subconcepts are given in Goh and Chua (2009).

An example of the proposed pruning is shown in Fig. 2 wherein the inverted tree structure in Fig. 2 represents the hazard identification tree that is being retrieved. The job step is a steel pipe cutting work that uses arc-welding equipment. The activity was executed in a temporary shed and near an excavator. For clarity sake, the consequences of each of the contact events are not presented in the hazard identification tree.

For illustration purposes, three events as highlighted in the figure are labeled for discussion. The indexes of Events 1, 2, and 3 are presented in the different call-out boxes. Events 1 and 2 are breakdown events (BE) and Event 3 is a contact event. During adaptation, the relevance of incident events in the hazard identification tree is ascertained by determining the LSS between indexes of each corresponding event with the input case's situational variables.

As illustrated, Event 1 is pruned off because its Resource index, "Resource=Arc-welding set," is dissimilar ($\mathrm{LSS}=0.1 < 0.6$) to the corresponding index of the input case, i.e., "Resource=Handheld rebar cutter." As reflected in Fig. 2, all events that follow Event 1 are also deleted. This is because each preceding event acts as a necessary condition for its following events. Thus once the preceding event is deleted, the subsequent events that are linked to it are also deleted.

On the other hand, the relevant Action and Object-worked-on indexes of Event 2, "Action=Cutting" ($\mathrm{LSS}=1$) and "Object-worked-on=Steel Pipe" ($\mathrm{LSS}=0.65$), are similar ($\mathrm{LSS} \geq 0.6$) to the corresponding situational variables of the input case, "Action=Cutting" and "Object-worked-on=Reinforcement Bar," respectively. Thus, Event 2 is deemed to be a possible incident event in the new situation and is not deleted. Accordingly, Event 2 of the adapted tree will also be updated to reflect the corresponding situational variables of the input case, namely "Action

J. Constr. Eng. Manage., 2010, 136(2): 170-178

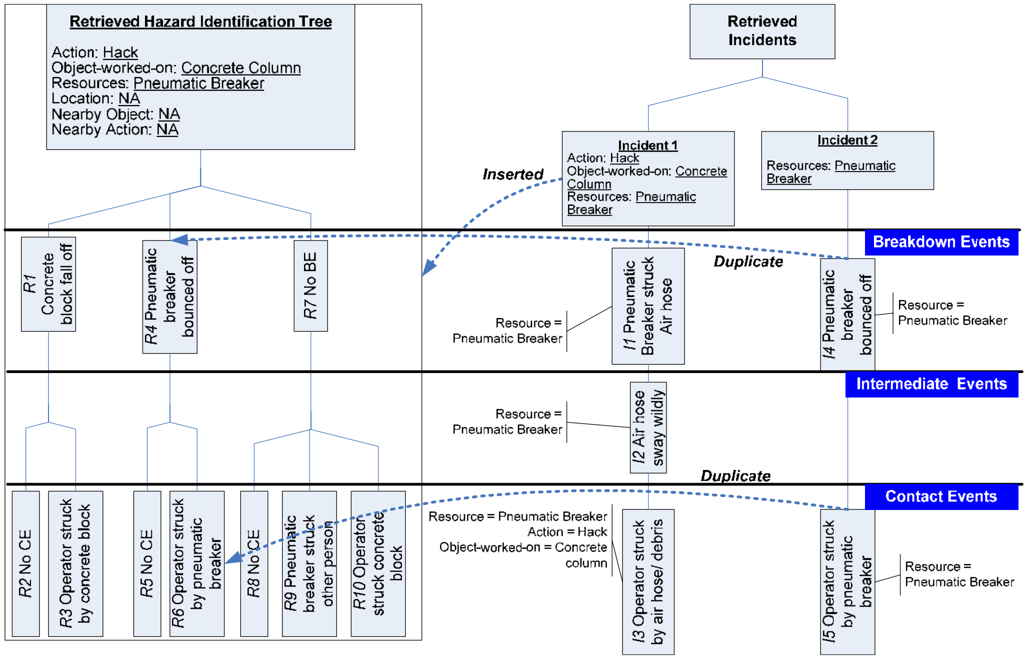

Fig. 3. Example of incident cases being integrated in to a hazard identification tree

=Cutting" and "Object-worked-on=Reinforcement Bar."

Similar to Event 1, Event 3 was deleted because its Resource index, "Resource=Arc-welding set," is dissimilar ($LSS=0.1 < 0.6$) to the corresponding index of the input case, i.e., "Resource=Handheld rebar cutter." The deletion was made despite its preceding event being accepted. This is because even though the preceding event is a necessary condition for Event 3 to be accepted, it is not a sufficient condition that guarantees its acceptance.

As can be seen from the above example, only events with indexes that are similar ($LSS \geq 0.6$) with the corresponding indexes of the input case, for instance Event 2, will be accepted. In the example, Events 1 and 3 only have one index and the events were deleted when their only index was dissimilar to the corresponding index of the input case. In the situation when an event has more than one index, the event will be deleted when any one of its indexes is dissimilar with the corresponding index of the input case. This is because indexes, like preceding events, are necessary conditions; whenever one of these necessary conditions is not met the event will be deleted.

Adaptation of Retrieved Incident Cases

The adaptation of retrieved incident cases is executed after the pruning of the retrieved hazard identification tree. It is performed in two parts. The incident cases are first pruned to remove irrelevant incident events. This is done in the same way as the pruning of the hazard identification trees through the LSS of the event nodes. Subsequently, incident events that are relevant but not identified earlier are inserted into the hazard identification tree.

The procedure for incorporating the incident cases into the hazard identification tree is outlined in Fig. 1 and illustrated with a hypothetical example depicted in Fig. 3. The left hand side of Fig. 3 shows a hazard identification tree that has been retrieved and pruned, and the situational variables depicted at the root node amended to reflect the case input. On the right hand side are two incident cases, incidents 1 and 2, which have been retrieved, based on similarity to the input case. Although hypothetical, Incidents 1 and 2 are based on actual incident cases obtained from the Land Transport Authority in Singapore (LTA).

All incident events of the two incidents have the index "Resource=Pneumatic breaker," while the contact event of Incident 1 has additional two indexes, "Action=Hack" and "Object-worked-on=Concrete column." Consequently, all the incident events are relevant to the input case ($LSS=1 \geq 0.6$ for all events), so that no pruning of incident events was necessary. In the case where the incident events have LSS less than 0.6, these incident events will be removed along with their child events.

During insertion phase, only those events that are not already contained in the hazard identification tree are inserted into the tree. An event will not be inserted if it occurs under the same parent event (or the root node) and has the same event value. In the case of incident 2 (refer to Fig. 3), breakdown event "Pneumatic breaker bounced off" ($I4$) is a duplicate (same as node R4 of the hazard identification tree) so that the event is not inserted. Subsequent child events will be checked in turn for duplication. If not duplicated, they will be inserted under the corresponding parent node of the hazard identification tree. In this example, the child event of $I4$, namely $I5$ is compared with the contact events of corresponding parent node $R4$ in the retrieved hazard identification tree. Since it is also duplicated, it is not inserted; otherwise it will be inserted as a new contact event under the corresponding parent node. However, duplicated events, such as $I5$, hold important frequency or likelihood data that can be used during risk analysis. The proposed CBR approach keeps track of these duplicated events for risk analysis purpose, but the discussion on risk analysis is beyond the scope of this paper.

Finally, to complete the adaptation, breakdown event $I1$ (Pneumatic breaker struck air hose) has no duplicated breakdown event node of the hazard identification tree. Consequently, it is inserted into the retrieved tree along with its child nodes $I2$ and $I3$. Thus, the final adapted hazard identification tree has been derived com-

J. Constr. Eng. Manage., 2010, 136(2): 170-178

Table 1. Case Titles of Hazard Identification Trees in Case Base

No.	Case title
1	Gas-cutting in confined space (tank)
2	Rigging up precast element
3	Lift precast wall using crawler crane
4	Arc welding of suspended pipes in trench
5	Concreting work using bucket
6	Lowering pipe into trench using excavator
7	Loading truck with soil using excavator
8	Concrete breaking
9	Frame scaffold erection
10	Gas-cutting of *H*-pile

prising the event nodes of the pruned hazard identification tree together with the event nodes of Incident 1.

Case Study

The case study described herein is an extension of the case study presented by Goh and Chua (2009) and it is meant to validate the adaptation mechanism discussed earlier. The case study utilizes actual data from the LTA and the Mine Safety and Health Administration (MSHA 2004).

The case base comprises ten hazard identification trees and 59 incident cases. The incident cases belong to a rapid rail project under LTA's charge and it involves the construction of precast viaducts and above ground train stations. The contract has 59 reported incidents after 1,444,8300 man-hours of work (project duration of three years and nine months). Most of the cases (66.1%) are less than three-man-days lost and were not required to be reported to the Ministry of Manpower in Singapore. Of the remaining 20 cases which were required by law to be reported to the Ministry of Manpower, there is a case of fatality in which one worker was killed. The incident cases occurred in different types of activity, such as soil boring, hoisting, concreting, and manual handling work. The cases also provided a variation of incident sequences thus supplying a rich source of knowledge for validation of the adaptation approach.

The case base contains ten hazard identification trees listed in Table 1. These hazard identification trees are based on risk assessments contained in safety documents obtained from different main contractors, LTA and Mine Safety and Health Administration (MSHA 2004). Two experienced construction safety practitioners (each with at least eight years in construction safety) were asked to verify the content of the cases and assign appropriate weights to assist in the assessment of the similarities between cases.

The job scenario used in the case study for validation is a common material delivery activity on site depicted in Fig. 4 showing a snapshot of the activity in which a lorry crane is in the process of unloading a bundle of timber strips. Based on the linguistic structure presented earlier, the situational variables for the job step, i.e., the input case, is "*Unloading (Action; Actn of Timber strip (Bundle) (Object-worked-on; OWO) using Chain sling (Resource; Res) and Lorry crane (Resource; Res) at Site entrance (Location; Loc) with nearby Plants/vehicles (nearby objects; NBO) and no nearby actions (NBA)*." Please see the Appendix on the abbreviations.

Fig. 4. Scenario for the hazard identification case study

Hazard Identification Adaptation

Fig. 5 shows the retrieved hazard identification tree being pruned (Step 1 of Fig. 1). The retrieved hazard identification tree is case 6 of Table 1 which involves the lowering of pipes into a trench using an excavator; the root node depicting the indexes of the retrieved tree. The indexes of the input case for lowering the timber strips are also shown for comparison. For each event node, the critical index with the lowest LSS is also noted for discussion. Altogether seven incident events were removed. The BEs, "Soil structure collapse" (Event 4) and "Soil/objects fall into trench/ excavation" (Event 1) were deleted because one of their indexes, "Location=Trench," is not similar to the Location index of the input case "Location=Site entrance" with LSS=0. Their child events were also deleted since they are conditional on their parent BEs. Contact event (Event 33), "Cut by lifted object" is also deleted because one of the event's indexes, "Object-worked-on (OWO)=Pipe" is dissimilar to the OWO (=timber strip (bundle) [index (b)] of the input case with LSS=0.11<0.6). The remaining incident events were not removed because their minimum LSS is at least 0.6 or higher.

Based on the GSS, five incident cases were retrieved as shown in Fig. 6. None of the incident cases need to be pruned (Step 2 of Fig. 1) because all the indexes of the events in the incidents match the indexes of the input case. The critical index with the minimum LSS are indicated for each node and none is less than the threshold value of 0.6. The incident cases are then checked for duplication with the incident sequences in the adapted hazard identification tree (Step 3 of Fig. 1). As indicated in Fig. 6, five of the incident events are duplicated. For example, breakdown event E1 "Lifting gear failure" of Case E is duplicated by event 23 of the adapted hazard identification tree of Fig. 5 and thus not inserted. The remaining five events have not been identified in the hazard identification tree before and are hence inserted. For example, breakdown event C1 "Runaway plant/vehicle" of Case C has not been identified before. The consequences of each of the

J. Constr. Eng. Manage., 2010, 136(2): 170-178

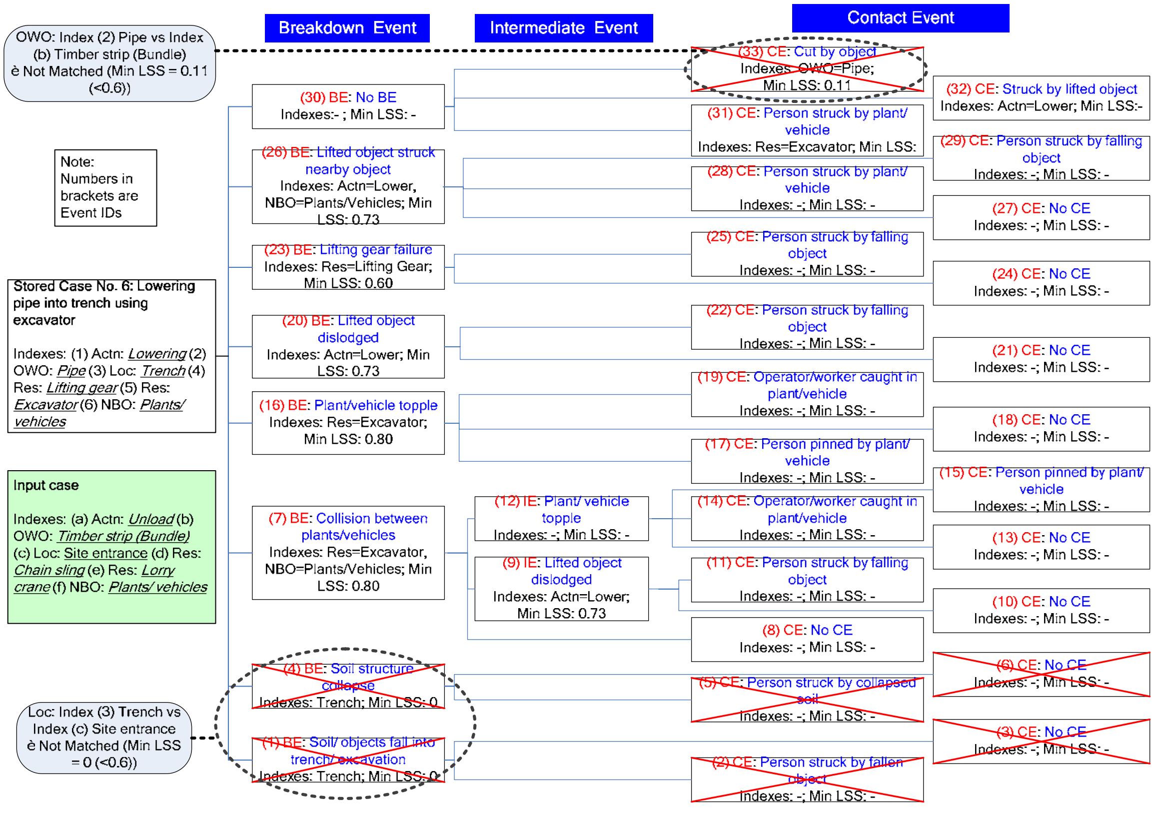

Fig. 5. Retrieved hazard identification tree being pruned

incident cases are used to estimate the risk of the input case quantitatively, but this is beyond the scope of this paper.

The final adapted hazard identification tree is shown in Fig. 7. The shaded incident events are based on the retrieved incident cases. These hazards have not been identified earlier but have only been found through analysis of relevant past cases and through adaptation they have been incorporated in current hazard identification tree. Note also that the situational variables at the root of the retrieved hazard identification tree have been adapted to the present case.

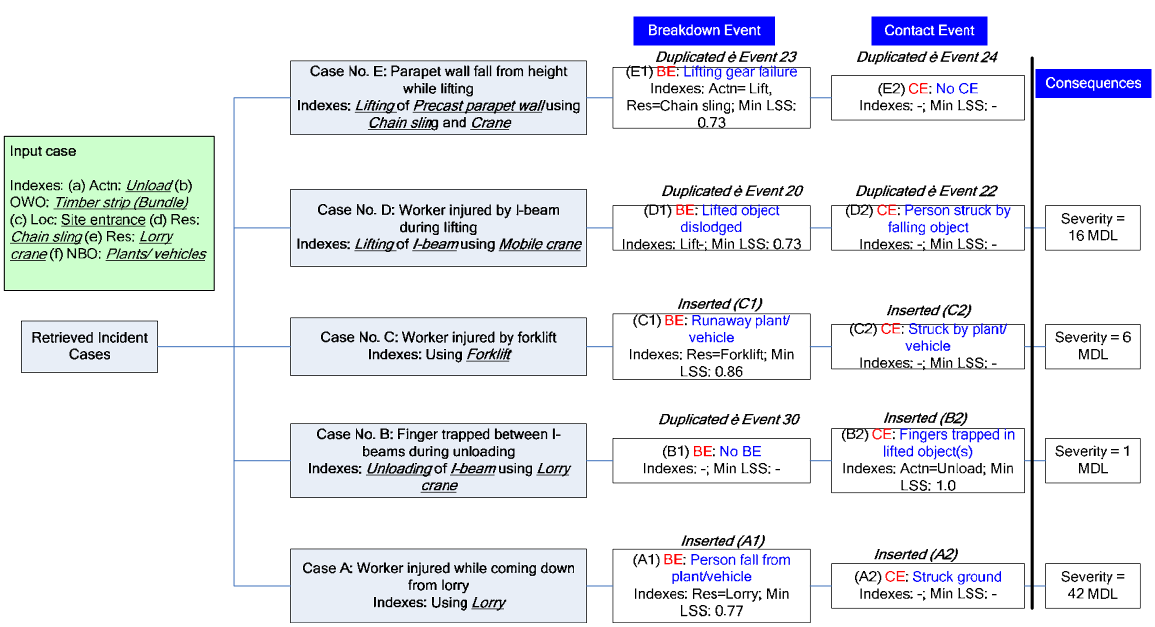

Fig. 6. Incident sequences of the retrieved incident case

J. Constr. Eng. Manage., 2010, 136(2): 170-178

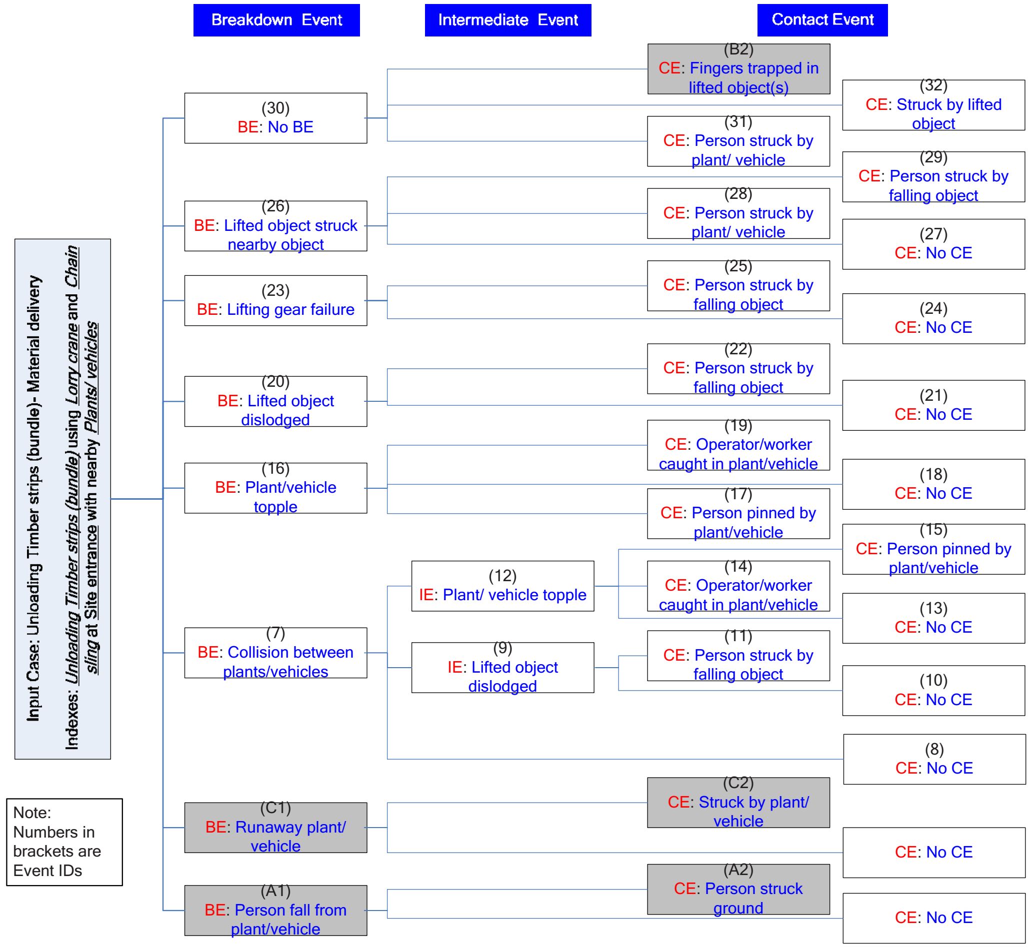

Fig. 7. Hazard identification tree after hazard identification adaptation

Discussions

The proposed CBR approach aims to facilitate feedback of past experiences so that hazard identification is more comprehensive and efficient. To achieve this aim, an automatic adaptation of retrieved past hazard identification tree and incident cases is very important. Without the adaptation process, the retrieved past cases may contain irrelevant portions that will be very cumbersome for hazard identification team to remove. As demonstrated in the case study, the hazard identification tree is automatically pruned to remove incident events deemed to be irrelevant. In addition, past incident cases were automatically integrated into the pruned hazard identification tree.

One of the benefits of the proposed approach is that the rationale for adaptation is provided based on the LSS. The similarity threshold can be modified by the user when necessary. For example, in the case study the contact event, "Cut by object," was deleted because the index of incident Event 33 (see Fig. 5), "Object-worked-on (OWO)=Pipe," has a low LSS (LSS=0.11) with the input case's situational variable, "OWO=Timber strip (bundle)." The user may still want to include the incident event because he might regard it as a possible incident event even if the OWO is a bundle of timber strips. Since each adaptation is based on the LSS, the user can easily review and verify the basis of hazard identification adaptation. Thus the LSS, which provides the reasoning for the adaptation, is useful for human interaction with the automated process.

The adaptation strategy ensures that relevant incident cases that are not already in the hazard identification tree are inserted. The benefit of this process is that the hazard identification teams will always be reminded of relevant incident occurrences, allowing measures to be implemented to prevent recurrence. Implementation of the approach will hence facilitate organizational learning, which is particularly problematic in the construction industry due to challenges in ensuring transfer of knowledge across projects.

As highlighted by Gadd et al. (2004), one of the key pitfalls in risk assessment is the use of generic risk assessment instead of site specific risk assessment. It is believed that most operational personnel in construction companies will be under certain amount of production pressure that will easily result in failure to adapt generic risk assessments to suit the site specific conditions. The proposed adaptation approach will help to improve the efficiency of the adaptation process by providing a draft site specific hazard identification, thereby reducing the occurrence of this common pitfall. However, it is emphasized that the proposed approach is not meant to replace human judgment, but to facilitate efficient recall of organizational experiences so as to improve the comprehensiveness and efficiency of hazard identification.

The proposed adaptation strategy and the overarching CBR approach is meant to improve hazard identification of a typical

J. Constr. Eng. Manage., 2010, 136(2): 170-178

risk management process in a construction project. For the approach to succeed, the CBR system, including the procedure for capturing and coding of incidents and hazard identifications, has to be designed as part of the construction safety risk management system. Therefore, the proposed approach is most viable in construction companies with established systems to conduct risk assessment, investigate and record incidents and document risk assessments. An alternative approach is to implement the system at industry level. A repository of hazard identification trees and incident cases can be created by an appropriate construction industry association through collection of data from its members. The association can then codify the past cases and allow its members to tap onto the repository during hazard identification. This approach has the advantage of having a comprehensive case base, which improves the range of hazards identified. Furthermore, having an association managing the repository makes it possible for smaller construction companies, who are usually weaker in hazard identification, to implement the proposed approach. However, the association will have to implement robust data management processes to prevent unintentional leakage of sensitive incident data.

The proposed CBR approach can be expanded to include risk analysis and risk control selection so that the efficiency and comprehensiveness of construction risk management can be improved further. Future research should explore how likelihood and severity data can be stored, retrieved and adapted to suit new risk analysis. The research on incorporating CBR concepts in construction risk analysis should identify ways to integrate both objective and subjective assessments of likelihood. Future research can modify the adaptation strategy to customize past risk controls to ensure their suitability to new work situations. In addition, other AI techniques like genetic algorithm and neural network can be evaluated to determine how a range of AI techniques can be combined to improve the quality of construction safety risk management.

Conclusions

Hazard identification is an important safety planning function. If not thoroughly implemented hazards will not be able to be eliminated or controlled. However due to work pressures or lack of experience foreseeable hazards may not be identified. This may occur despite actual incidents occurring before. This paper is a continuation of the work of Goh and Chua (2009), which discussed a proposed CBR approach to hazard identification. The adaptation approach is based on the concept of similarity of retrieved information and the situation at hand. The similarity scores calculated allow users to make amendments to the similarity threshold to facilitate specific needs. In addition, the structured approach facilitates implementation on a computer system. The proposed approach aims to utilize past knowledge in the form of past hazard identifications and incidents to ensure that construction companies are reminded of knowledge captured in these past cases.

Besides improving the quality of hazard identification in construction projects, the proposed adaptation mechanism facilitates efficient hazard identification. The proposed mechanism first prunes off less relevant incident events from the retrieved hazard identification tree. Subsequently, retrieved incident cases are pruned and then inserted into the pruned hazard identification tree to form the adapted hazard identification tree. The detailed review and adaptation of past cases are too cumbersome to be accomplished by hazard identification teams. Furthermore, traditional database retrieval would not have been able to conduct the adaptation of retrieved cases. The draft hazard identification produced by the proposed adaptation mechanism will reduce the effort and time of the hazard identification team significantly. As described in the paper, the approach had been validated with a case study.

The paper demonstrates that application of AI techniques in construction safety risk management is a feasible idea. Possible implementation strategies and future research agenda had been proposed.

Appendix. Common Abbreviations Used

- Actn—action.
- BE—breakdown event.
- CE—contact event.
- CSQ—consequences.
- IE—intermediate event.
- Loc—location.
- LSS—local similarity score.
- NBA—nearby actions.
- NBO—nearby objects.
- OWO—object-worked-on.
- Res—resource.

References

Balducelli, C., and D'Esposito, C. (2000). "Genetic agents in an EDSS system to optimize resources management and risk object evacuation." *Safety Sci.*, 35(1–3), 59–73.

Bergmann, R., and Wilke, W. (1998). "Towards a new formal model of transformational adaptation in case-based reasoning." *Sixth German Workshop on CBR*, L. Gierl and M. Lenz, eds., University of Rostock.

British Standards Institute. (2007). *Occupational health and safety management systems—Requirements*, British Standards Institute, London.

Chua, D. K. H., and Goh, Y. M. (2004). "An incident causation model for improving feedback of safety knowledge." *J. Constr. Eng. Manage.*, 130(4), 542–551.

Chua, D. K. H., Li, D. Z., and Chan, W. T. (2001). "Case-based reasoning approach in bid decision making." *J. Constr. Eng. Manage.*, 127(1), 35–45.

Empolis Knowledge Management GmbH. (2001). *CBR: Works 4-Compendium*.

Gadd, S. A., Keeley, D. M., and Balmforth, H. F. (2004). "Pitfalls in risk assessment: examples from the U.K." *Safety Sci.*, 42, 841–857.

Goh, Y. M., and Chua, D. K. H. (2009). "Case-based reasoning for construction hazard identification: Case representation and retrieval." *J. Constr. Eng. Manage.*, 135(11), 1181–1189.

Hillson, D., and Murray-Webster, R. (2007). *Understanding and managing risk attitude*, Aldershot, Burlington.

Kolodner, J. (1993). *Case-based reasoning*, Morgan Kaufmann, San Mateo, Calif.

Luu, D. T., Ng, S. T., and Chena, S. E. (2003). "A case-based procurement advisory system for construction." *Adv. Eng. Software*, 34, 429–438.

Mendes, J. R. P., Morooka, C. K., and Guilherme, I. R. (2003). "Case-based reasoning in offshore well design." *J. Pet. Sci. Eng.*, 40(1–2), 47–60.

Mount, C., and Liao, T. W. (2001). "Prototype of an intelligent failure analysis system." *ICCBR 2001, LNAI 2080*, D. W. Aha and I. Watson,

J. Constr. Eng. Manage., 2010, 136(2): 170-178

eds., Springer, Berlin Heidelberg, 716–730.
MSHA. (2004). "Mine safety and health administration—Training improves safety." ⟨http://www.msha.gov/training/part46/pt46train.htm⟩ (June 1, 2004).
Russell, S., and Norvig, P. (1995). *Artificial intelligence: A modern approach*, Prentice Hall, Englewood Cliffs, N.J.
Sinha, A. P., and May, J. H. (1996). "Providing design assistance: A case-based approach." *Information Systems Research*, 7(3), 363–387.
Suh, M. S., Jhee, W. C., Ko, Y. K., and Lee, A. (1998). "A case-based expert system approach for quality design." *Expert Sys. Applic.*, 15, 181–190.
Swartz, G. (2002). "Job hazard analysis—A primer on identifying and controlling hazards." *Prof. Saf.*, 27–33.

J. Constr. Eng. Manage., 2010, 136(2): 170-178