Generic self-learning decision support system for large-scale disasters

Stella Moehrle
Institute for Nuclear and Energy Technologies
Karlsruhe Institute of Technology, Germany
Stella.Moehrle@kit.edu

ABSTRACT

Large-scale disasters, particularly failures of critical infrastructures, are exceptional situations which cannot be solved with standard countermeasures. The crises are complex and the decision makers face acute time pressure to respond to the disaster. IT based decision support systems provide potential solutions and assist the decision making process. Many decision support systems in emergency response and management concentrate on one kind of disaster. Moreover, complex structures are modeled and recommendations are made rule-based. This work in progress paper describes the first steps towards the development of a generic and self-learning decision support system. The methodology used is case-based reasoning. The paper concludes with a sample emergency decision process.

Keywords

Decision support system, disaster management, case-based reasoning

INTRODUCTION

In the recent years not only the frequency of various crises increased but also the complexity grew. Especially failures of critical infrastructures have become of growing concern. Due to their complexity, they are highly vulnerable to disturbances, and especially in the field of critical infrastructure (CI) protection, there are a lot of methodologies and applications (Yusta, Correa and Lacal-Arántegui, 2011). Different modeling techniques are applied to CIs to describe the current state of the infrastructures and to understand the dynamic behavior using simulation techniques. Modeled systems and established rules run the risk of being too complex and therefore not being deployable in the event of disaster. Or modeled systems and established rules may be incomplete and therefore not able to capture non-standard situations. One approach of improvement is to use self-learning respectively case-based reasoning (CBR) systems to assist decision making in the event of a disaster. The idea is to imitate experts' behavior who can take advantages of their experiences and who learn over the course of time by gaining more experience. The core of the system is a knowledge database with data from previous disasters.

The objective of this paper is to illustrate the first steps in developing a decision support system (DSS) for complex crises with the focus on defining a general model and evolving a methodology to find potential solutions for emergency response. In the long run, the focus will be directed on the affected critical infrastructures. First, the CBR methodology will be explicated and some related work will be presented. Afterwards a general model will be formulated. Thereafter the parameters for the DSS will be specified including the case description attributes and the similarity function. This section concludes with a discussion concerning the reasoning step. Next, an application example is briefly described followed by a conclusion with a discussion regarding future work.

Case-based Reasoning

CBR solves new problems by utilizing knowledge of previously experienced problem situations. The methodology corresponds to a frequently applied way of human problem solving (Aamodt and Plaza, 1994). The fundamental assumptions are that similar problems have similar solutions and that solutions can be found more quickly by the confrontation with the same type of problem (Beierle and Kern-Isberner, 2008). The general procedure comprises identifying, assessing and describing the current problem situation (case), finding a similar case, reusing its solution, adapting and revising the solution, if necessary, to the current situation, evaluating the proposed solution and learning by storing the new case with its confirmed solution. The procedure is a cycle process which can be described by four steps or tasks: Retrieve, Reuse, Revise and Retain. Each step can be decomposed into sub-processes offering a range of different methods for using CBR. For a detailed discussion see Aamodt and Plaza, 1994.

Related Work

Otim, 2006 outlines fundamental concepts of a case-based knowledge management system for disaster management and discusses the advantages. For example, it can be used in domains that do not have underlying models. Further, CBR is a model of human reasoning. Therefore we can believe more easily the validity of the received solutions of the system than of complex models. Relevant information from previous disasters is gathered prior to the disaster and decision makers can make use of some kind of precedents which accelerates the decision making process. Nextcase/safety (Virkki-Hatakka and Reniers, 2009) and the Fire Emergency Handling Recommender (FEHR) (Chakraborty, Ghosh, Maji, Garnaik and Debnath, 2010) are inspiring application examples. The first serves as a support for safety managers deciding on prevention measures. To determine the main characteristics of the accidents descriptive statistics was used. The adaptation step is based on fuzzy calculations. As the model output values of this paper are not numerical but measures in the event of disaster the adaptation step cannot be calculated easily. Moreover, the characteristics of cases are investigated by experts because using statistical methods require a large database of past events. FEHR serves as a support for fire fighters or administrators concerning required resources to handle a fire. The cases have ontological representation. As in the DSS of this paper FEHR has been implemented by jColibri2. In FEHR, the model parameters are configured by the user. The adaptation step is just concerned with numerical values by using direct proportions.

MODEL

One of the central questions is how to present knowledge that is to say how to illustrate past events and their solutions. Bergmann, Kolodner and Plaza (2005) give a brief overview of the main types of case representation. A case is subdivided into a problem/solution description. As the aim is to develop a generic DSS the case representation has to cover different kinds of disasters. One approach is to describe past events as a vector of attribute-value pairs. The attribute ranges are numeric, symbol, temporal and textual. When using symbolic types it is possible to provide an enumeration of the allowed values which can, for example, have a flat structure or be arranged in an ontology. There are several approaches to define the attributes to describe a disaster. For example, the German Federal Office of Civil Protection and Disaster Assistance provides damage parameters for risk analysis. They could be used to represent disasters concerning their loss and damages. Another example is DesInventar, a methodology and tool for building disaster databases and analyzing them. DesInventar provides a disaster reporting format to store disasters in a structured way. Particularly it provides a classification of events, attributes to describe the effects of disasters and predefined vocabulary to specify the type of cause. However, besides the structure for disaster representation, a unique value range for symbolic attributes is just as important to be defined. Moreover, a classification of the allowed values leads to a more profound knowledge representation. In many cases, either no value ranges are defined for symbolic values or just flat structures for them are provided. The choice of employed attributes for the DSS is inspired by Tactical situation object (TSO) – a message structure for disaster and emergency management. The idea behind is to support the transfer of information between computer-based systems by encoding disaster/emergency relevant terms in an XML Schema. TSO provides the relevant attributes to describe a disaster or emergency and it defines a basic vocabulary for disaster management with unique expressions. The codes are arranged hierarchically which provides a basic categorization of event features. The solution consists of a bundle of measures which are temporarily ordered. Measures have requirements and impacts and therefore they cannot be implemented arbitrarily. In the following the general definition of the model proposed is presented. Most of the definitions are based on the definitions in Stahl, 2004.

An ***attribute*** A is a pair (A_{name}, A_{range}) where A_{name} is a unique label and A_{range} is the value range. Further, $a_{name} \in A_{range} \cup \{\text{ NULL }\}$ denotes the current value of A_{name}. A ***case description model*** CD is a finite ordered list of attributes $CD = (A_1, \ldots, A_n)$ with $n > 0$. Let $\overline{A} := A_{1_{range}} \times \ldots \times A_{n_{range}}$ and $\overline{S}$ be the space of possible solutions for large-scale disasters. A ***solution bundle*** S is an ordered list of solutions

$S = (s_1, \ldots, s_m)$ with $s_i \in \bar{S}, i = 1,\ldots,m, m \geq 0$. A ***case model*** C is a pair of two finite ordered lists $C = (CD, S) = ((A_1, \ldots, A_n), (s_1, \ldots, s_m)) \in \bar{A} \times \bar{S} \times \ldots \times \bar{S}$. A ***case*** according to a given case model C is a pair $c = (cd, s)$ with $cd = (a_1, \ldots, a_n)$, $a_i \in A_{i_{range}} \cup \{\text{NULL}\}, i = 1, \ldots, n$ and $s = (s_1, \ldots, s_m)$, $s_i \in \bar{S}$, $i = 1, \ldots, m$. cd is the case description of C and s is the appendant solution bundle. A ***query*** is a case with an empty solution part.

In the following let q, c be two cases with $q = (q_1, \ldots, q_n)$ and $c = (c_1, \ldots, c_n)$. A ***local similarity measure*** for an attribute A is a function $S_A : A_{range} \times A_{range} \rightarrow [0,1]$. An ***attribute weight vector*** for a case description model is a vector $w = (w_1, \ldots, w_n)$ with $w_i \in [0,1]$ and $\sum w_i = 1$. Each w_i is called an attribute weight for A_i. A ***global similarity measure*** for a case description model is a function $S_G : \bar{A} \times \bar{A} \rightarrow [0,1]$ with $S_G(q,c) = f(S_1(q_1,c_1), \ldots, S_n(q_n,c_n), w)$ where S_i is the local similarity function for A_i, $i = 1, \ldots, n$ and w is the weight vector. The function $f : [0,1]^{2n} \rightarrow [0,1]$ is called the ***aggregation function***.

Attributes

Attribute	Attribute name	Attribute description	Value range
A_1	caseType	Scenario leading to the event	See TSO codes; terms are stored in an ontology for the case category. The ontology essentially comprises the codes including some additionally concepts. The idea is to refine the hierarchy levels.
A_2	actor	Endangered objects	Same principle as for caseType
A_3	locationType	Location where the event is taking place	Same principle as for caseType
A_4	environment	Context of the event	Same principle as for caseType
A_5	scale	Severity of the event	Levels 1-5 which reflect resources necessary to manage the disaster
A_6	riskAssessment	Evolution of the event	Increasing, stable or decreasing
A_7	affectedPersons	Affected Persons	Number of affected persons
A_8	triage	Triage categories	Number of persons triaged red, yellow, green, black
A_9	geo	Position type and geo coordinates	Position type: point, circle, surface
A_{10}	cause	Cause of event	Accidental, deliberate, natural
A_{11}	description	Textual description	Text

Table 1. Attributes of the case description model

Similarity Function

For the first four attributes a cosine similarity measure is proposed, each (Recio-García, Díaz-Agudo, González-Calero and Sánchez-Ruiz, 2006). The more concepts the individuals share the more similar are they. The values of the attributes 'scale' and 'riskAssessment' are compared by considering their distance in relation to their value ranges. The values of the latter one are first transformed into numerical values. To compare the number of affected persons the difference is considered and compared to the maximum and minimum of all differences concerning the affected people in the case base (difference-based similarity function, Stahl, 2004). The aim is to position the difference between the number of affected persons in the context of the differences of the numbers of affected persons of the case base. The values of the triage categories are written as a vector. An euclidean based similarity function normalized to the interval [0,1] is applied. The comparison of the geographical locations is carried out by contrasting the areas of the respective affected regions. The similarity measure of the attribute 'cause' is a simple String comparison. In order to test how similar textual descriptions of events are Apache Lucene is used which is a full-featured search engine library. The search mechanism is based on a statistical approach. The global similarity function computes the average of the local similarities.

The Reasoning Step

The solution of a case is a bundle of countermeasures with a fixed sequence. Instead of considering the most similar case from the past and adapting the solution bundle to the current circumstances the most similar cases are regarded. Assume that the decision makers set a similarity threshold of θ. Then the solution bundles considered arise from cases which similarities to the query are greater or equal than θ. Now, the decision maker can choose appropriate countermeasures from the different solutions bundles recommended to form a new solution bundle.

SIMPLE APPLICATION EXAMPLE

The case base consists of sixteen events gathered from the news and translated into TSO codes. The cases consist of rail accidents involving hazardous and non-hazardous goods, fire disasters, chemical accidents, transport accidents and power failure. The query is, for example, a deliberate release of cyanuric chloride of a chemical plant unless the energy supply can be restored within three hours and 22 minutes. The immediate environment especially the residents are endangered. The three most similar cases are the following:

Case	Measures	Similarity
Chemical accident, release of ammonia gas	- Residents have to close the windows - Water curtain in wind direction	62,4%
Chemical accident, release of nitrogen tetroxide because of a leakage over the valve connection of the pressure vessel	- Partial evacuation of the plant - Water curtain in wind direction - Traffic closure - Immission measuring	60,6%
Chemical accident, release of cyanuric chloride, leakage over a pneumatic conveyor line	- Stop of the conveyor line - Internal and external restricted zones - Water curtain in wind direction - Traffic closure	59,3%

Table 2. Results of the sample query

The decision maker can now choose the appropriate countermeasures and generate a new solution bundle. The revision also includes the specification of additional measures. This decision-making process speeds up the implementation of measures in a complex crisis as it presents potential solutions to solve the crisis. In this example the three most similar cases are chemical accidents. An interesting question is if and how sub-events of different disaster types can be compared in order to gain appropriate countermeasures. This will be further examined in connection with researches on the attribute weights. They influence the retrieving step significantly.

CONCLUSION AND FUTURE WORK

This paper presents the first steps in developing a generic decision support system for large-scale disasters. The methodology chosen is case-based reasoning. The focus is on the formal definition of the model and the description of the events. For the case description TSO is used. The nearest neighbor scoring is chosen for the retrieval method. Due to the different value ranges of the attributes local similarity functions are defined. The solution of a case is a bundle of countermeasures which are logically ordered. The reasoning step consists of merging the recommended solution bundles appropriately. This step includes deleting improperly measures and supplementing necessary ones. Following the model definition a small application example is shown.

For future work the case model will be enhanced by lessons that can be learned from past events. Moreover, the case description attributes have to be revised. The attributes used are too general to describe an event deeply.

Further, additional information is needed to learn about causalities in disaster events. When defining the global similarity function, one crucial factor is the attribute weight vector. One approach is to define disaster-type dependent weight factors which will be supported by experts. Moreover it will be examined if and how disasters of different disaster types can be compared concerning particular sub-events in order to gain some appropriate countermeasures. The reasoning step will be automated by adapting the recommended solution to the current situation. Instead of applying rules to adapt one solution bundle to a new situation, several solution bundles are used to find an adequate sequence of countermeasures. The idea is to store countermeasures in a graph and to learn pre-conditions and dependencies of countermeasures over the course of time. In the case retrieval step the relevant sub-graphs are localized to find an appropriate solution for the current event. The combination of self-learning methodology and graphs is especially suitable for complex crises. The domain has not to be completely modeled in advance. Decisions are not completely made rule-based and complex interdependencies can be represented well in graphs. In the long run, the DSS aims to focus on countermeasures concerning critical infrastructures. The objective of the DSS is to recommend countermeasures in the event of the disaster. Moreover, the system should give early warnings about potentially endangered objects and critical infrastructures, respectively, along with the appropriate countermeasures to avoid further damage. The next steps focus on earthquakes, flood water and extreme weather events to extend the set of case description attributes. Afterwards, the appropriate weight factors are determined. The work is supported by experts for the different disaster types.

ACKNOWLEDGMENTS

This work is supported by the Center for Disaster Management and Risk Reduction Technology (CEDIM), an interdisciplinary research center in the field of disaster management founded by Helmholtz Centre Potsdam-German Research Centre for Geoscience (GFZ) and Karlsruhe Institute of Technology (KIT).

REFERENCES

1. Aamodt, A. and Plaza, E. (1994) Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, *AI Communications*, 7, 1, 39-59
2. Beierle, C. and Kern-Isberner, G. (2008) Methoden wissensbasierter Systeme: Grundlagen, Algorithmen, Anwendungen, Vieweg + Teubner
3. Bergmann, R., Kolodner, J. and Plaza, E. (2006) Representation in case-based reasoning, *The Knowledge Engineering Review*, 20, 3, 209-213
4. Chakraborty, B., Ghosh, D., Maji, R. K., Garnaik, S. and Debnath, N. (2010) Knowledge Management with Case-Based Reasoning Applied on Fire Emergency Handling, *Industrial Informatics (INDIN), 2010 8th IEEE International Conference on*, 708-713
5. Otim, S. (2006) A Case-Based Knowledge Management System for Disaster Management: Fundamental Concepts, *Proceedings of the 3rd International ISCRAM Conference (B. Van de Walle and M. Turoff, eds.),* Newark, NJ (USA)
6. Recio-García, J. A., Díaz-Agudo, B., González-Calero, P. A. and Sánchez-Ruiz, A. A. (2006) Ontology based CBR with jCOLIBRI, *Applications and Innovations in Intelligent Systems XIV. Proceedings of AI-2006, the Twenty-sixth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, Springer, Cambridge, United Kingdom
7. Stahl, A. (2004) Learning of Knowledge-Intensive Similarity Measures in Case-Based Reasoning, PhD-Thesis, Technische Universität Kaiserslautern, dissertation.de
8. Virkki-Hatakka, T. and Reniers, G. L. L. (2009) A case-based reasoning safety decision-support tool: Nextcase/safety, *Expert Systems with Applications*, 36, 7, 10374-10380
9. Yusta, J. M., Correa, G. J. and Lacal-Arántegui, R. (2011) Methodologies and applications for critical infrastructure protection: State-of-the-art, *Energy Policy*, 39, 10, 6100-6119
10. DesInventar, available at http://www.desinventar.org (last accessed on 14 Jan 2012)
11. jColibri2, available at http://gaia.fdi.ucm.es/research/colibri/jcolibri (last accessed on 14 Jan 2012)
12. German Federal Office of Civil Protection and Disaster Assistance (2010) Methode für die Risikoanalyse im Bevölkerungsschutz
13. Tactical Situation Object (TSO), available at http://www.tacticalsituationobject.org (last accessed on 14 Jan 2012)