Using Ontologies to Support Model-based Exploration of the Dependencies between Causes and Consequences of Hazards

Abigail Parisaca Vargas and Robin Bloomfield
Centre for Software Reliability, City University, London, U.K.

Keywords: Preliminary Hazard Analysis, Hazard Analysis, Hazard Identification, Ontologies, Railway.

Abstract: Hazard identification and hazard analysis are difficult and essential parts of safety engineering. These activities are very demanding and mostly manual. There is an increasing need for improved analysis tools and techniques. In this paper we report research that focuses on supporting the early stages of hazard identification. A state-based hazard analysis process is presented to explore dependencies between causes and consequences of hazards. The process can be used to automate the analysis of preliminary hazard worksheets with the aims of making them more precise, disambiguating causal relationships, and supporting the proper definition of system boundaries. An application example is presented for a railway system.

1 INTRODUCTION

Hazard identification and hazard analysis are difficult and essential parts of safety engineering. These activities are very demanding and mostly manual. There is an increasing need for improved analysis tools and techniques. In this paper we report research that focuses on supporting the early stages of hazard identification. Hazard analysis is the identification of hazards and their initiating causes (International Organization for Standardization, 2000). It can be thought of as the process of investigating an accident before it actually occurs. Its aim is to exhaustively identify all possible causes of accidents, so that they can be eliminated or controlled before they occur (Leveson, 2011). Hazard analysis is extremely important, and lack of completeness in the analysis can have serious consequences.

A hazard is a potential source of harm. The term includes danger to persons arising within a short time scale, for example, fire and explosion, and also those that have a long-term effect on a person's health, such as release of a toxic substance (ISO, 1999). Harm is a physical injury or damage to the health of people or damage to property or the environment (ISO, 1999).

Hazard identification (HazID) is one of the most important parts of the hazard analysis, as it forms the basis of the activities carried out to design a safe system. By looking at the system from a safety perspective, safety analysts check important aspects of the system and try to identify hazards that could have been accidentally overlooked. Common methods for identifying hazards at the early stages of system design are checklists, HAZOPs (Hazard and Operability Studies) and PHA (Preliminary Hazard Analysis) (Kletz, 2001).

Various studies asserted that the most significant flaws in hazard analysis techniques applied at the early stages of system design are often related to the omission of hazards and hazard causes (Hardy, 2010). One reason for this could be that most hazard analysis techniques rely on the manual analysis of information recorded in text format using different spreadsheet representations, called hazard identification worksheets. The most basic description of a worksheet would be a table of three columns, where the first column describes a hazard, the second column describes the cause or causes of that hazard, and the third column describes the consequences of the hazard occurring. Different examples of this basic description will be discussed further in this paper.

Data and facts are usually abundant in hazard analysis worksheets. Because of this, for complex systems, there is always a good chance that even expert analysts could accidentally miss something, and therefore draw conclusions on the basis of incomplete information. A mechanised process for systematic exploration of the dependencies between hazards, causes and consequences of hazards would therefore be a convenient way of supporting experts during the analysis.

Various initiatives in the safety sectors are exploring the use of ontologies as a means to mechanise

Vargas, A. and Bloomfield, R..
Using Ontologies to Support Model-based Exploration of the Dependencies between Causes and Consequences of Hazards.
In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015) - Volume 2: KEOD*, pages 316-327
ISBN: 978-989-758-158-8

the analysis of causal dependencies in hazard analysis techniques. The aims of these initiatives are generally to capture consensual knowledge, reuse it and share it across different application domains and different teams (Corcho et al., 2003). For example, in (Mayer et al., 2008), an ontology is described for representing risk and risk assessment processes, as well as concepts such as hazards and threats. Others have explored the use of ontologies to support standard hazard analysis techniques, such as Job Hazard Analysis (e.g., see (Wang and Boukamp, 2009)), HAZOP (e.g., see (Stralhane et al., 2010; Daramola et al., 2011)), and FMEA (e.g., see (LEE, 2001)), or other hazard analysis techniques targeted to industrial sectors like construction, food supply (Letia and Groza, 2010; Yang et al., 2012), geology (Liu et al., 2010), and many others.

Despite all the examples above, little attention has been devoted to using ontologies for the analysis of preliminary hazard worksheets generated at the early stages of system design. Being able to analyse these preliminary worksheets is highly desirable, as issues can be resolved earlier in the development process, when it is easier and cheaper to fix problems. The present work aims to address this gap.

Contribution. In this paper we present an approach to facilitate the analysis of preliminary hazard worksheets generated at the early stages of system design. The approach facilitates the resolution of weaknesses in the hazard identification process. In our approach, an ontology is used to review preliminary hazard worksheets using a mechanised process based on a set of inference rules. Our approach allows one to check well-formedness of hazards, their causes, and consequences, as well as to discover new relationships between hazards, causes and consequences, if these were accidentally omitted in the hazard worksheets. The final result obtained using our approach is therefore an improved version of preliminary hazard worksheets, with disambiguated hazards and polished causal relationships.

Organisation. The paper is organised as follows. In Section 2, we contrast and compare our work with related work. In Section 3, we introduce our ontology. Then, in Section 4, we explain our proposed method, its steps, and the reasoning we used. In Section 5, we illustrate the method on a case study. Finally, in Section 6, we briefly conclude, summarise the results, and discuss future work.

2 RELATED WORK

A number of different hazard analysis techniques have been created over the last fifty years, and they are currently widely used by safety-critical industries. There are different examples of their use in complex systems (Leach, 2010; Zhang et al., 2010; Center for Devices and Radiological Health, US FDA, 2010). There are also examples of adaptations of standard hazard analysis techniques for identifying security hazards (Winther et al., 2001).

Despite the wide use of the standard hazard analysis techniques mentioned above, there is also strong criticism of them. New techniques are, in fact, being proposed that are specifically designed for the analysis of hazards in today's complex socio-technical systems. For example, Nancy Leveson describes a new approach to hazard analysis, STPA (System-Theoretic Process Analysis) (Leveson, 2011), based on the STAMP causality model. Another example is the Ontological Hazard Analysis (OHA) (Ladkin, 2005; Ladkin, 2010) proposed by Peter Ladkin for the analysis and maintenance of safety hazard lists using a refinement approach.

Differently from the efforts cited above, we do not aim to introduce a novel hazard analysis technique. Rather, we propose a mechanised method that can improve the analysis of results obtained at the early stages of system design. In the method presented in this paper, we start from preliminary hazard worksheets, and use an ontology to check consistency and well-formedness of information contained in these preliminary documents.

Finally, in his book called HAZOP and HAZAN (Kletz, 2001), Trevor Kletz discusses the feasibility of the automation of HAZOPs, for example, applying techniques from Artificial Intelligence, and whether these techniques could replace the safety analyst. Kletz concludes it is impossible. He gives two important objections to the automation of HAZOPs:

1. Artificial intelligence techniques can manipulate logical rules, but logic is just one aspect of human intelligence. For example, most of the scientists who have recounted how they came to make an important discovery or to achieve a significant breakthrough have stressed that when they found the answer to the crucial problem they intuitively recognised it to be right, and only subsequently went back and worked out why it was right (Kletz, 2001).
2. Knowledge used in HAZOPs is broad and deep, while expert systems are suitable only for narrow and deep knowledge (Kletz, 2001).

Logic is not able to represent all the knowledge and expertise of the safety analyst. We do not aim

to represent this; instead, we want to represent the knowledge already captured by hazard identification techniques, usually in worksheets. We want to be able to find logical relationships, which could have not been seen, and point out the relevant information that might have been hidden because of the quantity of information represented.

3 DEVELOPING THE ONTOLOGY PROTOTYPE

Ontologies, or explicit representations of domain concepts, provide the basic structure or framework around which knowledge bases can be built (Devedzić, 2002; Swartout and Tate, 1999). Ontologies are specific, high-level models of knowledge underlying all things, concepts, and phenomena. As with other models, ontologies do not represent the entire world of interest. Rather, ontology designers select aspects of reality relevant to their tasks (Devedzić, 2002; Valente et al., 1999).

During the design of the ontology prototype, we used a UML-like modelling language. It has been argued that UML can be used to model ontologies (Cranefield and Purvis, 1999), yet there are recent, more specialised modelling languages, like OntoUML (Benevides et al., 2010).

We decided to use the ontology engineering environment Protégé because it provides an integrated environment for the ontology development and it has different features supporting different tasks during the ontology life cycle. The knowledge representation language of choice was OWL-DL because it is the most used by the ontology development community (Simperl et al., 2009), and because OWL-DL and SWRL offer a number of sophisticated reasoning capabilities.

3.1 Ontology Domains

When developing an ontology, we aim to formally and explicitly describe concepts in a domain, as well as their properties. Describing the concepts involves defining classes for these concepts and arranging them in a hierarchy.

The first domain we work on is the HazID worksheets HI, which consist of different hazards with multiple causes and where different consequences might exist.

Table 1 shows a basic example of the main part of a HazID worksheet. The example is taken from (Evans and Associates, 2006). The hazard description for hazards $A-5$ is *Toxic gases in tunnel*.

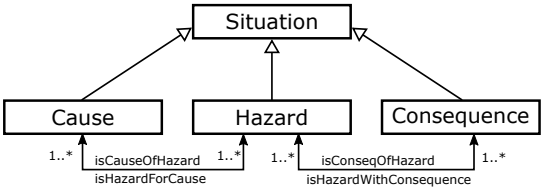

Figure 1: Examples of concepts that are part of our ontology.

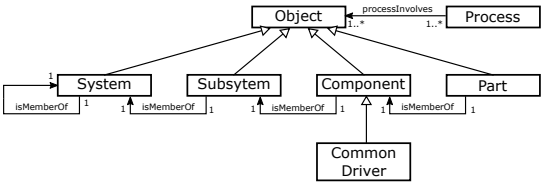

Figure 2: Part of Ontology.

The cause of this hazard is *Toxic gases enter in tunnel from alignment or station*. The alignment in the context of a railway is the ground plan of the railway, as well as the path that the train follows. The consequences for hazard $A-5$ are *injury*, *death* or *service disruption*. Another cause of hazard *Toxic gases in tunnel* is *Maintenance personnel release toxic gas while performing work*.

Figure 1 shows the concepts of Hazard, Cause, and Consequence represented in our ontology and the relationships among them. Boxes in the figure represent the concepts, arrows with triangular heads represent inheritance, and arrows with two heads represent the relationships. The concepts can be used to define a set of hazards, H, consisting of $h_1 \dots h_n$ individual hazards; a set of causes, consisting of $c_1 \dots c_n$ individual causes; a set of consequences, consisting of $q_1 \dots q_n$ individual consequences. Figure 1 shows the relationships among these concepts. The relationships cause-hazard and hazard-consequence are many-to-many. That is, a cause might occur on different hazards, and a consequence might occur on several hazards. In the hierarchy of types, Situation is a higher node and the super type for Hazard, Cause and Consequence.

Hazard identification is performed within a system and its boundaries. In order to represent our basic ontology, we need to define two core concepts to represent the characteristics of a system and its behaviour. We chose Object Process Methodology (OPM) (Crawley and Dori, 2011) because it is geared towards modelling systems in general. From an ontological perspective, OPM's building blocks are objects and processes. We use these concepts in order to represent the initial design of the system and its behaviour. An object can be a system, subsystem, component or a part of the system. The object can be

Table 1: Basic example of HazID worksheet.

Hazard	*Cause*	*Consequence*
$A-5$	Toxic gases enter in tunnel from alignment or station	Injury, death,
Toxic gases in tunnel	Maintenance personnel release toxic gas while performing work	service disruption

in a particular state. Processes transform objects by generating, consuming, or changing their state (Crawley and Dori, 2011). A process can be decomposed into sub-processes. Figure 7 shows an example of OPM's building blocks, where *Transporting* is a process. This process involves the objects *Tracks*, *Alignment*, *LRV*, *Platform* and *Station*.

"System", is an important concept we need to represent in our ontology. The physical domain of the system is composed of subsystems. Subsystems, in turn, are composed of components, and components are composed of parts. From the CLIOS formalism (Sussman et al., 2009), we take the notion of *Common drivers*, which are shared components among subsystems of a system.

Figure 2 shows the concepts of Object, Process, System, Subsystem, Component, Common Driver and Part Modelled. The initial design of a system S, consists of different natural divisions such as subsystem B, component M, part P. We identify a set of subsystems, B, consisting of $b_1 \dots b_n$ individual subsystems; a set of components, M, consisting of $m_1 \dots m_n$ individual components; and finally, a set of parts, P, consisting of individual parts $p_1 \dots p_n$. A part might occur within a component, so then the relationship is one-to-one. Similarly, the relationship between component and subsystem and between subsystem and system is one-to-one. In addition, a system has a recursive relationship, where a system can be part of another system, and the relationship will again be one-to-one.

The behavioural description of the system is described as the set of objects, O, consisting of $o_1 \dots o_n$ individual objects; the set of processes, C, consisting of $c_1 \dots c_n$ individual processes. An object may participate in different processes and a process may occur in different objects, so the relationship is many-to-many.

Figure 3 shows that subsystem A is a member of system B. They have different system boundaries and, as a result, concentrate on different hazards. System B provides the environment for subsystem A. A hazard, normally, is described with respect to its system boundary, so it is a relative term.

Figure 5 shows fundamental concepts of the basic ontology. "Thing" is a generalisation of objects, processes and situations; it represents the class of all things, or the abstract objects that can be described by the criteria for being something. This concept is

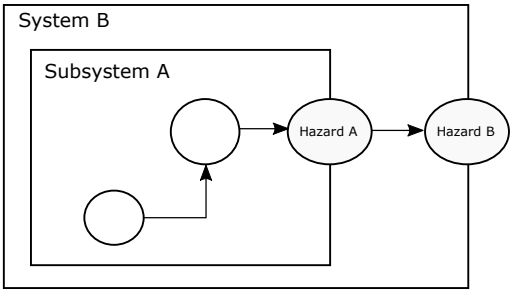

Figure 3: Hazard in Systems. Based on (Railtrack, 2001).

taken from class-defined ontologies. In addition, Figure 6 shows four new relationships that we need to discuss. The hazard information within the ontology needs to include information such as the scope and location of the hazards. When we talk about the scope of a hazard, we refer to what affects the hazard directly; we therefore refer to two kind of scopes: *behavioralScope* and *physicalScope*. If a hazard affects a determined system, that system will be the scope of the hazard. When we refer to the *physicalScope* we refer to the boundary of the hazard. For example, if subsystem A is a member of system B, then system B is the scope of subsystem A. When we talk about location, we refer to where the hazard is located in terms of objects and also in terms of processes; the relationships that we will use will be called *physicalLocation* and *behaviouralLocation*.

Hazards, causes and consequences could be interlinked by any of these characteristics. This kind of knowledge is already known by the safety analyst or it can be deduced by reading the HazID description. The idea is that information provided by these relationships (object properties in Protégé) offer a natural way to establish associations among hazard meaning, or senses, in a certain HazID process. This can be profitably used during the processing of the hazard worksheets, e.g., to disambiguate the process. For example, the hazard *virus affecting the system*, when referring to an infusion pump medical device, is ambiguous between its Biology and Computer Science senses and, therefore, the system boundary of the hazard, can be disambiguated by assigning the correct domains to the contexts where it actually occurs.

3.2 Extended Ontology

In this subsection we explain in detail the extended

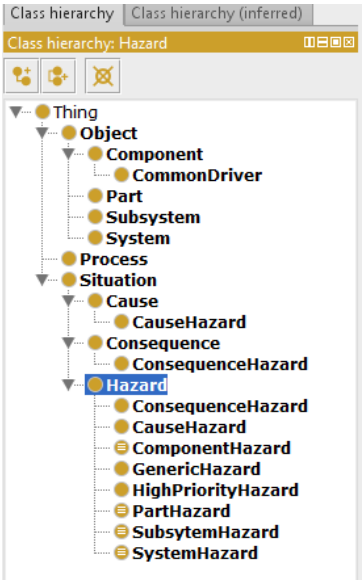

Figure 4: Class hierarchy of complete ontology.

ontology. The extended ontology is formed using more specific types, i.e., lower nodes in the hierarchy. This provides a more granular organisation of the hazards, causes and consequences, and relationships between them. Figure 4 shows a more detailed hierarchy of the classes in our ontology.

Hazards have a scope. In order to generate a hierarchy to hazards according to their scope, we create equivalent classes called *PartHazard*, *ComponentHazard*, *SubsystemHazard* and *SystemHazard*. When editing the ontology, the analyst can decide which of the *SystemHazards* belong to the *HighPriorityHazards*. These are the hazards that are deemed the most important for analysis by the safety analyst. The class *GenericHazards* is useful to handle worksheets with generic hazards identified using, e.g., a Preliminary Hazard Analysis.

The class *ConsequenceHazard* is a subclass of class *Hazard*. The classes *Consequence* and *CauseHazard* are subclasses of *Hazards* and *Cause*.

4 THE METHOD

We developed an ontology to capture domain information related to hazards, systems, and hazards within a system. The ontology is designed to help the analysts in the task of hazard identification using the knowledge already gathered, and the relationships that already exist. An ontology provides the means to structure information by classes, property descriptions and relationships between classes and individuals. Analysts can use the ontology for reasoning

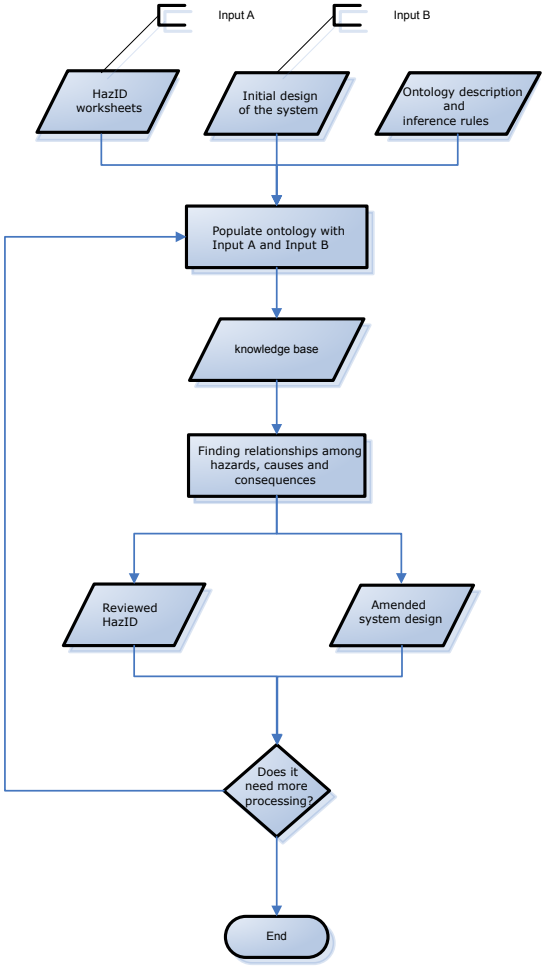

Figure 5: Control flow diagram of our method.

about missing relationships between hazards, causes and consequences. Figure 5 shows the steps of our method:

1. We start from an initial Preliminary Hazard Analysis, an initial design of the system, an ontology description of the domains, and a set of rules for reasoning about our domains. The ontology description models the data from the inputs illustrated above, including property descriptions and relationships between classes and individuals. We used the open-source editor and framework Protégé 5.0 (Stanford Center for Biomedical Informatics Research, 2015). The output is our knowledge base.
2. The next step is to find possible indirect causal and overlooked relationships. This is ilustrated in Sections 5.1 and 5.2. The intention is to use the structured and automated reasoning provided by tools for the analysis of ontologies, such as Protégé's built-in reasoner, Pellet.

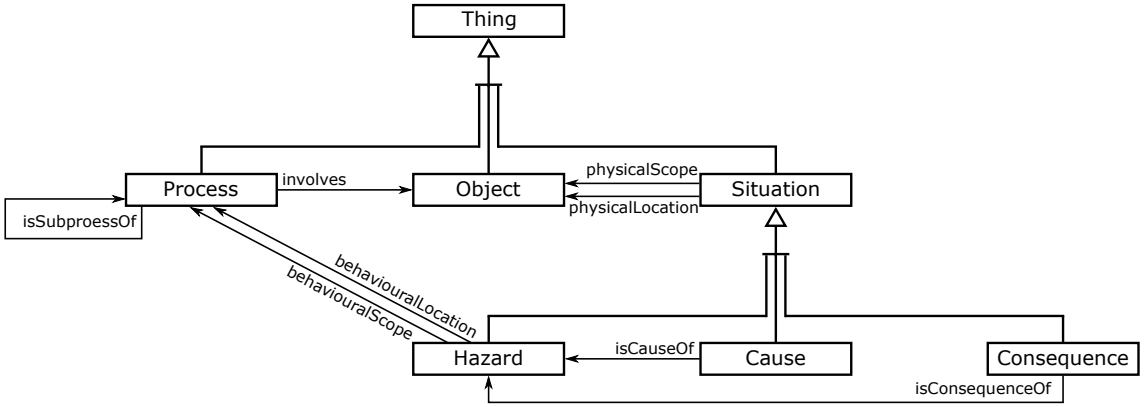

Figure 6: Core classes of the ontology.

3. The initial results are used to review the HazID worksheet, and perhaps modify the System Design model. Yet more could be done after the first amendment, and our knowledge base could still be reviewed and amended further.

4.1 Reasoning using the Ontology

In this subsection we will explain some concepts that are needed to understand the reasoning approach implemented with the ontology. The approach will also be illustrated with concrete examples.

4.1.1 Direct and Indirect Causality

In (Leveson, 2011) Leveson argues that current hazard analysis techniques are event chain models of the form: "if event *A* had not occurred, then the following event *B* would not have occurred". This type of model supports limited notions of direct causality, and it is sometimes impossible to incorporate indirect relationships. For example, consider the following statement: "Smoking causes lung cancer". It is argued (Leveson, 2011) that in current models such a statement would not be allowed; however, it is widely accepted that there is a relationship between the two, even though the relationship is complex and indirect.

4.1.2 Synonymy, Entailment and Non-direct Causality

Hazard identification is a demanding task and relies on detailed descriptions. It is difficult to remember each and every hazard and cause in a hazard analysis and, as a consequence, the worksheets may refer to hazards, causes and consequences in different written forms. Sometimes this reflects real differences that should be captured, while in other situations the differences are only artificial, and should therefore be amended. For example, when analysing a road system, we could have the following hazards: $H-1$ *Vehicle drives on the wrong lane*; $H-2$ *Vehicle passes through the wrong lane*; $H-3$ *Vehicle on the wrong lane*. $H-1$ and $H-3$ could have the same meaning, or $H-3$ could mean that the vehicle actually stopped on the wrong lane. $H-1$ and $H-2$ could also have the same meaning if each just means that a vehicle temporarily drives on the wrong lane. An ontology could help the analyst to find and identify these hazards, as well as to disambiguate them, when needed, by means of logical rules applied to the structured knowledge.

When two situations (hazards, causes, consequences) have the same meaning, we will call them synonyms. Using only inference rules, we cannot determine whether a hazard is a synonym of another hazard, cause or consequence. To resolve the problem, we rely on different relationships established between situations that are part of the HazID entry. For example, two synonym hazards might have the same physical location, physical scope, behavioural location, or behavioural scope. Therefore, using rules to find hazards where multiple similar relationships occur can help to spot synonyms. In addition, a synonym is not the only kind of relationship that can be spotted. For example, let us consider hazard $H-4$ *Train on fire*. The cause of this hazard is $Ca-1$ *Train stopped adjacent to a wayside fire*. A train circulates around different places in a railway system; it also can stop at different places. For example, it can stop at the platform, which can be on fire. Following this reasoning, there might be a relationship between a hazard called $H-5$ *Fire on station platform* and $Ca-1$ *Train stopped adjacent to a wayside fire*. Therefore, there might be a nonlinear (indirect) causal relationship between $H-5$ *Fire on station platform* and $H-4$ *Train on fire*.

In logic, entailment (Fellbaum, 1990), or strict

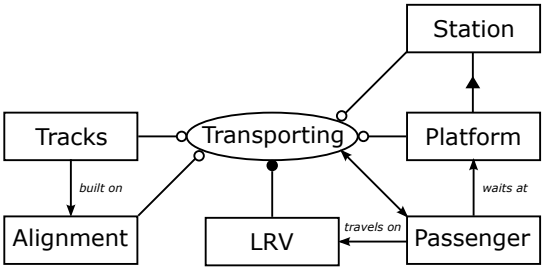

Figure 7: Top level representation of our initial system design in OPM.

implication, is properly defined for propositions; a proposition P entails a proposition Q if and only if there is no conceivable state of affairs that could make P true and Q false. Entailment is a semantic relationship because it involves reference to the states of affairs that P and Q represent. We generalise the term in order to talk about hazards. For us, in the context of the hazard descriptions for a certain system, entailment refers to the special relationship between two hazards, where a hazard, $H1$, of a more general hazard, $H2$, also entails $H2$. In addition, if $H1$ entails $H2$, then it can not be the case that $H2$ entails $H1$. In order to explain this, we introduce a brief example. We mentioned before that the alignment in the context of a railway is the ground plan of the railway, as well as the path that the train follows. The crossing is the crossroad between the train path and the motor road. Hazards $H-6$ *Motor vehicle on alignment* and $H-7$ *Road vehicle drives around crossing gate* refer to a motor vehicle inside the train path. Because the crossing is part of the path of the train, $H-7$ entails $H-6$, because while a motor vehicle drives around the crossing gate, the motor vehicle is on the alignment. In addition, $H-6$ does not entail $H-7$ because *Motor vehicle on alignment* does not necessarily mean that the vehicle is driving around the alignment. Using inference rules and a reasoner could therefore help the safety analyst to disambiguate the meaning of those hazards and also to find hazards interlinked by the causation relationships. The following section will exemplify some of these rules and their application using a case study based on a railway system.

5 CASE STUDY

We now introduce a case study that we use to illustrate the method: the West Corridor Light Rail Transit (LRT) System. Figure 7 shows a representation of the system in Object Process Modelling (OPM). The hazard analysis that we use is the Preliminary Hazard Analysis (PHA) for the West Corridor LRT Project Design Phase (Evans and Associates, 2006). The description of the system was inferred from the PHA document and from terminology in rail transport and railways (www.allenrailroad.com, 2014; Transport Canada, 2014; www.railway-technical.com, 2014; www.trafficsigns.us, 2014; www.railsigns.uk, 2014).

We succinctly describe the system:

- Alignment provides a path and physical infrastructure, such as tunnels and crossrails.
- Track is the structure consisting of the rails, fasteners, sleepers, and ballast, plus the underlying subgrade. The track provides the physical structure by which the Light Rail Vehicle (LRV) circulates.
- A station is a railway facility where trains regularly stop to load or unload passengers and/or freight.
- A control system provides help in controlling the train to prevent accidents and to improve circulation. The service provided can be concisely described as: train separation or collision avoidance, line speed enforcement, enforcing temporary speed restrictions, enabling rail worker wayside safety.
- An Overhead Catenary System (OCS) comprises different components, such as wires suspended between poles, bridges supporting overhead contact wires which are normally energised with electricity, etc. The OCS powers the LRV.
- The LRV is the vehicle that circulates along the railway and transports passengers.

Figure 7 represents only the top level of the process representation. The process Transporting is related to the objects by the relationship processInvolves. In this case, the process involves the objects LRV, Alignment, Tunnel, Track and Platform. Subprocesses of Transporting are: Carrying passengers, Circulating, Grade crossing management, Guiding LRV, Loading and unloading passengers, and Transmitting power. All these processes are related to objects and have subprocesses as well. We will mention some of the processes while developing our examples in the sections below.

5.1 Example with Non-direct Causality

Let us start from a small subset of the PHA performed. Below, we define the rule that helps us to find non-direct or indirect causality between hazards and causes. The property representing this relationship is *isIndirectCauseOf*.

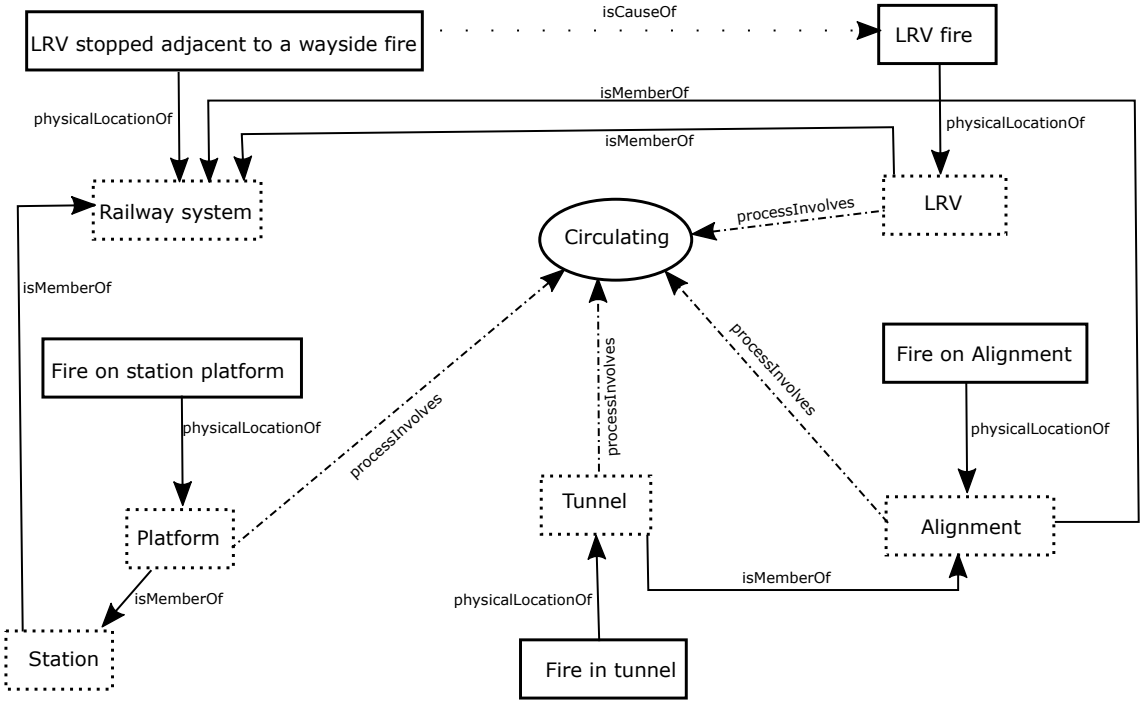

Figure 8: Relationships among individuals, exemplifying the relationships that are used in example 5.1.

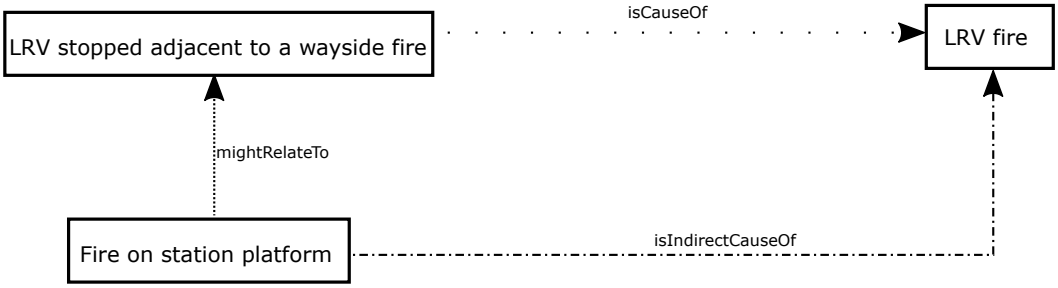

Figure 9: Example of non-direct casualty.

$$
\begin{aligned}
&\mathit{physicalLocationOf}(?q,?o),\\
&\mathit{behaviouralLocationOf}(?p,?a),\\
&\mathit{Object}(?o),\\
&\mathit{Situation}(?p),\\
&\mathit{processInvolves}(?a,?o),\\
&\mathit{physicalLocationOf}(?p,?o1),\\
&\mathit{isMemberOf}(?o,?o1),\\
&\mathit{Situation}(?q),\\
&\mathit{Process}(?a)\\
&\rightarrow \mathit{mightRelate}(?p,?q)
\end{aligned}
\tag{1}
$$

This rule is explained as follows. Every situation happens within some object (based on the boundary of hazards). The property representing this is $\mathit{physicalLocationOf}(a,b)$. We need to represent this property since any situation, that could affect the behaviour of the system, depends on where this situation occurs. In our ontology, the behaviour of the system is described via processes; a process may involve various objects ($\mathit{processInvolves}(e,f)$ property). We need also to record which process(es) could be directly affected by a situation, and we do this using the property $\mathit{behaviouralLocationOf}(c,d)$.

Figure 8 helps us to understand better how it is going to happen. We can see three kinds of things (Objects, Situations and Processes) and different relationships among them, and that all are related by the process *Circulating*. This process describes the actions of the LRV passing directly through different objects. Therefore, the *Circulating* process involves objects such as *tunnel*, *alignment*, *platform* at a station and railway *crossings*. The situation *LRV stopped adjacent to a wayside fire* refers to anywhere within the *Railway system*. This situation affects the process

Table 2: Table representing a small subset of the PHA for the case study (Evans and Associates, 2006).

Hazard	*Cause*
$A-1$ LRV fire	$Ca-1$ LRV stopped adjacent to a wayside fire
$A-2$ Fire/smoke on station platform	$Ca-2$ Electrical wiring fault
$A-3$ Fire/smoke on alignment	$Ca-3$ Fire on station $Ca-4$ Fire on wayside building or brush
$A-4$ Fire/smoke on station platform	$Ca-4$ Electrical wiring fault

Circulating. Also, there exists a relationship among the situations *Fire on station platform*, *Fire in tunnel* and*Fire on alignment* because the physical locations of all these situations are involved with the process *Circulating*. Last, the physical locations of all those different situations are related by a transitive, hierarchical relationship *MemberOf*.

Rule 1, in our example, returns as an answer that Cause *LRV stopped adjacent to a wayside fire* might relate to (*mightRelate*) hazards *Fire on station platform*, *Fire on tunnel*, *Fire on Alignment* and *LRV fire*. It is not a reflexive relationship. *LRV stopped adjacent to a wayside fire* is definitely related to *LRV fire* because the former is a direct cause of the latter. What we could conclude is that any of those hazards could be indirect causes of *LRV fire*. Figure 9 illustrates this, and we can add this relationship (*isIndirectCauseOf*) to our knowledge base.

5.2 Another Example on Non-direct Causality

Let us examine another example. We will first give some definitions related to railways. A grade crossing is an intersection where a road passes across a line of railway at a grade. A grade is a part of a railway or road that slopes upwards or downwards. A grade crossing warning system consist of all of the elements or warning devices such as signals, signs, lights and horns that automatically alert the public that a train is approaching. Figure 10 illustrates that *Grade crossing warning system* is part of *Crossing in a railway*. This object has a clear functionality, which is process *Warning at crossing*. Situations $TC-6$ and $Ca-8$ both happened within object *Grade crossing warning system* and refer to the behaviour of process *Warning at crossing*. In addition, the physical scope of both situations is the wider object *Crossing* (*Grade crossing warning system* is part of *Crossing*). Rule 2, below, helps us to discover this possible relationship. The property *physicalScope* refers to the boundary of the hazard. As explained earlier, if a hazard affects a determined system, that system will be the scope of the hazard. The property *isSubProcessOf* exemplifies that a process can be decomposed into subprocesses. The rest of the properties have been explained while describing Rule 1.

$$\begin{aligned} & \textit{physicalLocationOf}(?h, ?l2), \\ & \textit{behaviouralLocationOf}(?s, ?p1), \\ & \textit{behaviouralLocationOf}(?h, ?p2), \\ & \textit{physicalScopeOf}(?h, ?c), \\ & \textit{physicalLocationOf}(?s, ?l1), \\ & \textit{physicalScopeOf}(?s, ?c), \\ & \textit{isSubProcessOf}(?p1, ?p3), \\ & \textit{Situation}(?s), \\ & \textit{isSubProcessOf}(?p2, ?p3), \\ & \textit{Situation}(?h) \\ & \rightarrow \textit{possibleEntailment}(?s, ?h) \end{aligned} \tag{2}$$

Using Rule 2, we can discover a relationships between cause *Insufficient warning before gate descends* and hazard *Grade crossing warning system failed, or not visible or audible*. They might not be synonyms but there is a possible *entailment* relation between them. The final decision about this relationship is left to the analyst. What we can conclude is that there is a causality relationship between hazard *Grade crossing warning system failed, or not visible or audible* and hazard *Road vehicle breaks gate arm and fouls track* as shown in Figure 11. In the next section we draw the conclusions and discuss future work.

6 DISCUSSION AND CONCLUSION

In this paper we defined a state-based hazard process for model-based exploration of the dependencies between causes and consequences of hazards. In addition, we showed how ontologies can be used to reason about these dependencies. This method can be used to automate the analysis of preliminary hazard worksheets with the aims of making them more precise, disambiguating causal relationships, and supporting the proper definition of system boundaries. Our analysis process is supported by an ontology that can help analysts to find nonlinear causal relationships, and to disambiguate identified hazards and causes. A

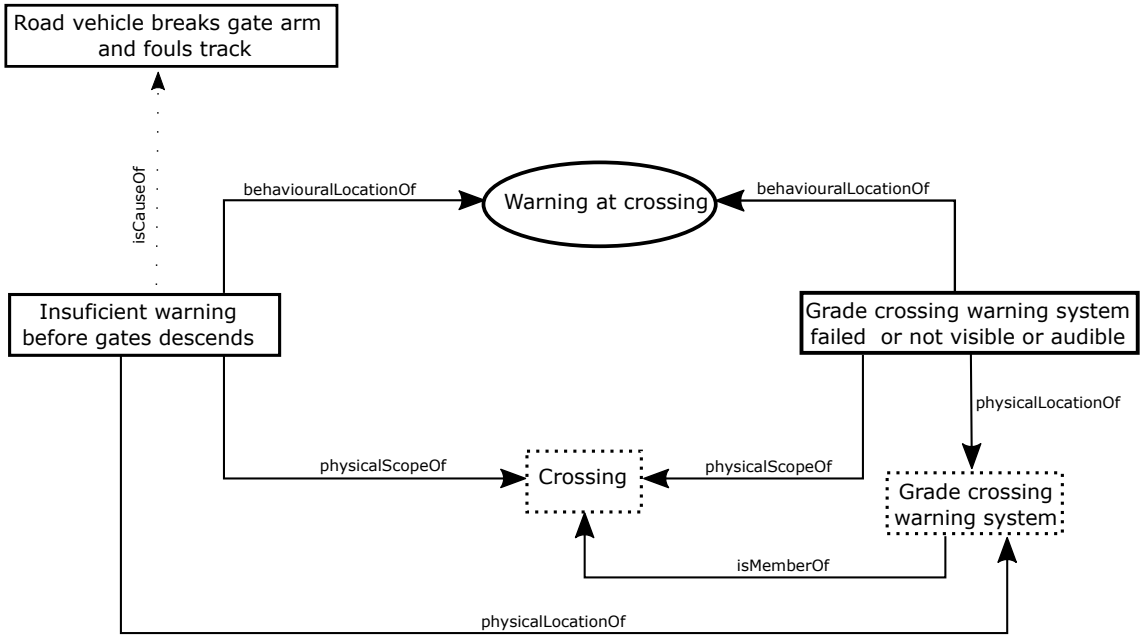

Figure 10: Relationships among individuals, exemplifying the relations that are used in example 5.2.

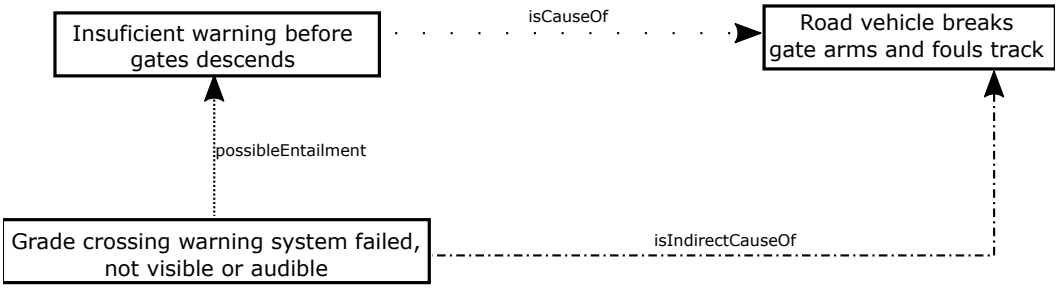

Figure 11: Possible causality relationship.

Table 3: Table that represents another small subset of the PHA for the case study (Evans and Associates, 2006).

Hazard	*Cause*
$TC-8$ Road vehicle breaks gate arm and fouls track	$Ca-8$ Insufficient warning before gate descends
$TC-6$ Grade crossing warning system failed, or not visible or audible	$Ca-6$ Equipment malfunction

demonstration of the approach was given using an industrial example, the PHA of the West Corridor LRT System (Evans and Associates, 2006).

A thorough validation of our ontology is in progress. We are currently working on two examples related to medical infusion pumps and an e-voting system. This will allow us to instantiate the ontology for real world situations, and check that the ontology is truly capturing the domain conceptualisation. This is a well grounded evaluation process (de Almeida Falbo, 2014).

Specialised tools are needed to better support our analysis process. We believe that for a safety analyst using Protege 5.0 might add more work to their tasks. Our idea is that there should be a front-end tool that interacts with the analyst and with the ontology. This way, the input of data will be more user-friendly to the safety analyst and the results could also be presented in a more understandable way to the user.

REFERENCES

Benevides, A. B., Guizzardi, G., Braga, B. F. B., and Almeida, J. P. A. (2010). Validating modal aspects of ontouml conceptual models using automatically generated visual world structures. *J. UCS*, 16(20):2904–2933.

Center for Devices and Radiological Health, US FDA (2010). Total product life cycle: infusion pump – premarket notification [510(k)] submissions.

Corcho, O., Fernández-López, M., and Gómez-Pérez, A. (2003). Methodologies, tools and languages for building ontologies: Where is their meeting point? *Data Knowl. Eng.*, 46(1):41–64.

Cranefield, S. and Purvis, M. (1999). Uml as an ontology modelling language. In *In Proceedings of the Workshop on Intelligent Information Integration, 16th International Joint Conference on Artificial Intelligence (IJCAI-99*, pages 46–53.

Crawley, E. and Dori, D. (2011). *Object-Process Methodology: A Holistic Systems Paradigm*. Springer Berlin Heidelberg.

Daramola, O., Stralhane, T., Sindre, G., and Omoronyia, I. (2011). Enabling hazard identification from requirements and reuse-oriented hazop analysis. In *Managing Requirements Knowledge (MARK), 2011 Fourth International Workshop on*.

de Almeida Falbo, R. (2014). Sabio: Systematic approach for building ontologies. In *Proceedings of the 1st Joint Workshop ONTO.COM / ODISE on Ontologies in Conceptual Modeling and Information Systems Engineering co-located with 8th International Conference on Formal Ontology in Information Systems, ONTO.COM/ODISE@FOIS 2014, Rio de Janeiro, Brazil, September 21, 2014.*

Devedzić, V. (2002). Understanding ontological engineering. *Commun. ACM*, 45(4):136–144.

Evans, D. and Associates, I. (2006). West corridor LRT project final engineering design phase preliminary hazard analysis draft revision 0. http://www.rtd-fastracks.com/media/uploads/wc/WC_Risk_Assessment_Draft_06-06.pdf.

Fellbaum, C. (1990). English verbs as a semantic net. *International Journal of Lexicography*, 3(4):278–301.

Hardy, T. (2010). Using accident reports to improve the hazard identification process. *28th International System Safety Conference*.

International Organization for Standardization (2000). *ISO 14971: medical devices - application of risk management to medical devices*. ISO.

ISO (1999). *ISO/IEC GUIDE 51:1999(E): Safety Aspects - Guidelines for Their Inclusion in Standards*. Australian Standards.

Kletz, T. (2001). *Hazop and Hazan: Identifying and Assessing Process Industry Hazards*. The Institution of Chemical Engineers.

Ladkin, P. B. (2005). Ontological analysis. *Safety Systems*, 14(3).

Ladkin, P. B. (2010). Ontological hazard analysis of a communication bus system. Technical report, Causalis Limited.

Leach, K. (2010). *A Demonstration of the Applicability of HAZOP to Voting Systems*. New Castle University.

LEE, B. H. (2001). Using fmea models and ontologies to build diagnostic models. *AI EDAM*, 15:281–293.

Letia, I. and Groza, A. (2010). Developing hazard ontology for supporting haccp systems in food supply chains. In *Intelligent Systems and Informatics (SISY), 2010 8th International Symposium on*, pages 57–62.

Leveson, N. G. (2011). *Engineering a safer world: Systems thinking applied to safety*. MIT Press.

Liu, G., Wang, Y., and Wu, C. (2010). Research and application of geological hazard domain ontology. In *Geoinformatics, 2010 18th International Conference on*, pages 1–6.

Mayer, R., Plank, C., Bohner, A., Kollarits, S., Corsini, A., Ronchetti, F., Siegel, H., Noessing, L., Mair, V., Sulzenbacher, U., Tosoni, D., Cimarosto, S., Zanco, A., Todorov, S., Krastev, L., Wergles, N., Gasperl, W., Mayerl, M., Toli, T., Haradalia, H., Koutsias, N., Kreuzer, S., Liehr, C., Rachoy, C., Papez, J., and Jindra, P. (2008). Monitor: Hazard monitoring for risk assessment and risk communication. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 2(4):195–222.

Railtrack (2001). *Engineering Safety Management*, volume 1 and 2. Railtrack.

Simperl, E., Mochol, M., Bürger, T., and Popov, I. O. (2009). Achieving maturity: The state of practice in ontology engineering in 2009. In *Proceedings of the Confederated International Conferences, CoopIS, DOA, IS, and ODBASE 2009 on On the Move to Meaningful Internet Systems: Part II*, OTM '09, pages 983–991, Berlin, Heidelberg. Springer-Verlag.

Stanford Center for Biomedical Informatics Research, S. U. S. o. M. (2015). Protege 5.0 beta version. http://protege.stanford.edu/.

Stralhane, T., Omoronyia, I., and Reicenbach, F. (2010). Ontology-guided requirements and safety analysis. In *6th International Conference on Safety of Industrial Automated Systems*.

Sussman, J., Dodder, R. S., McConnel, J. B., Mostashari, A., and Sgouridis, S. (2009). The clios process a user's guide. https://esd.mit.edu/Faculty_Pages/sussman/CLIOS-PROCESS.pdf.

Swartout, W. and Tate, A. (1999). Guest editors' introduction: Ontologies. *IEEE Intelligent Systems*, 14(1):18–19.

Transport Canada (2014). Rail transportation. http://www.tc.gc.ca/eng/rail-menu.htm.

Valente, A., Russ, T., MacGregor, R., and Swartout, W. (1999). Building and (re)using an ontology of air campaign planning. *IEEE Intelligent Systems*, 14(1):27–36.

Wang, H.-H. and Boukamp, F. (2009). *Ontology-Based Job Hazard Analysis Support*, chapter 66, pages 676–685. American Society of Civil Engineers.

Winther, R., Johnsen, O.-A., and Gran, B. (2001). Security assessments of safety critical systems using hazops. In *Computer Safety, Reliability and Security*, volume 2187 of *Lecture Notes in Computer Science*, pages 14–24. Springer-Verlag.

www.allenrailroad.com (2014). Railroad glossary and definitions. http://www.allenrailroad.com/consulting/Railroad_Glossary.htm.

www.railsigns.uk (2014). Rail signs and signals of great britain. http://www.railsigns.uk.

www.railway-technical.com (2014). Railway technical web pages. http://www.railway-technical.com.

www.trafficsigns.us (2014). Manual of traffic signs. http://www.trafficsign.us/railsign.html.

Yang, X., Gao, R., Han, Z., and Sui, X. (2012). Ontology-based hazard information extraction from chinese food complaint documents. In Tan, Y., Shi, Y., and Ji, Z., editors, *Advances in Swarm Intelligence*, volume 7332 of *Lecture Notes in Computer Science*, pages 155–163. Springer Berlin Heidelberg.

Zhang, Y., Jones, P. L., and Jetley, R. (2010). A hazard analysis for a generic insulin infusion pump. *Journal of diabetes science and technology*, 4(2):263.