Safety Science 69 (2014) 3–17

Contents lists available at SciVerse ScienceDirect

Safety Science

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

Safety management – A multi-level control problem ☆

Björn Wahlström [a,*], Carl Rollenhagen [b]

[a] Vattenfall SMI, AX-22920 Brändö, Finland
[b] Department of Philosophy, Royal Institute of Technology, SE-100 44 Stockholm, Sweden

ARTICLE INFO

Article history:
Available online 4 July 2013

Keywords:
Safety management
Control structures
Modelling safety
Methods and tools
Nuclear power

ABSTRACT

Activities in safety management build on a control metaphor by which control loops are built into the man, technology, organisational and information (MTOI) systems to ensure a continued safety of the operated systems. In this paper we take a closer look on concepts of control theory to investigate their relationships with safety management. We argue that successful control relies on four necessary conditions, i.e. a system model, observability, controllability and a preference function. The control metaphor suggests a division of the state space of the modelled system into regions of safe and unsafe states. Models created for selected subsystems of the MTOI-system provide a focus for control design and safety assessments. Limitations in predicting system response place impediments to risk assessments, which suggest that new complementary approaches would be needed. We propose that polycentric control may provide a concept to consider in a search for a path forward. We investigate approaches for modelling management systems and safety management. In spite of promises in the use of a control metaphor for safety management there are still dilemmas that have to be solved case by case. As a conclusion we argue that the control metaphor provides useful insights in suggesting requirements on and designs of safety management systems. The paper draws on experience from the Vattenfall Safety Management Institute (SMI), which started its operation in 2006.

1. Introduction

Members of societies are faced with risks emanating from both old and new technological systems. Accidents that have occurred in nuclear power plants, chemical production facilities, off-shore installations, and many other industrial branches, present a growing concern in society. Focus on technological factors, human factors, and the more recent conceptual innovation of safety culture, has provided safety managers with new tools and methods for safety management. Moreover, the safety paradigm has shifted from focus on parts towards a holistic conceptualisation of safety, but exactly what this means in terms of practical safety management is not always clear. One difficulty seems to be that the systems involved encompass very diverse areas of knowledge and modelling approaches, making it hard to find a common language for describing different aspects of system behaviour.

Many authors have argued that systemic approaches to safety should be applied (Bang Dyhrberg and Langaa Jensen, 2004; Kirwan, 2011). We generally support such approaches. In the present contribution we argue that a systemic approach entails an understanding of the specific characteristics that govern the behaviour of generic subsystems found in socio-technical systems. Already the need for considering people and organisations in system models calls for inputs from areas such as anthropology, psychology, social psychology, sociology, management science and economics. In a search for a common approach for modelling systems, subsystems and their interactions, we argue that a control metaphor may provide an overarching language that can be used both in design and analysis of safety management.

Applying a control metaphor for understanding system behaviour is not new. For example, it has been applied at several hierarchical levels for understanding safety of large complex systems by Rasmussen and Svedung (2000), Kjellén (2000) and Swuste et al. (2010), among many others. To illustrate, the society exercises control of companies that operate hazardous facilities through laws and regulations. Organisations implement policies and management systems to ensure that plants are designed, built, operated and maintained in a safe manner. Controls are implemented in many processes and they include both feedback and feed forward paths by which outcomes are monitored and correcting actions are initiated when deficiencies are found. Our contribution rests on the assumption that in order to efficiently use a control metaphor to support safety management, it is necessary to, at least in gross terms, identify a set of generic component classes in terms of man (people), technology, organisation and information (Rollenhagen, 2003).

☆ This is an extended version of a paper that was presented on the PSAM11 conference in Helsinki, 25–29 June, 2012.

* Corresponding author. Tel.: +358400448710; fax: +35898030781.
E-mail address: bjorn@bewas.fi (B. Wahlström).

0925-7535/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ssci.2013.06.002

The aim of this paper is to establish a basis for applying the control metaphor to create a frame of modelling, which can be used in analysis and design of systems and their controls to ensure that an acceptable safety can be reached. We are here mainly interested in process safety (Grote, 2012) for complex interconnected systems such as nuclear power plants, but the suggested approach can also be used for other hazardous systems. The advantage is that system models can be built using an approach that enables a consistent transfer of focus from the whole system toward increasing details. In this way it is possible to model various parts of a production system with their controls to provide evidence that the whole system can be operated safely.

If we can manage safety of a system with respect both to entirety and to details, it would be possible to define *necessary* conditions for safety, which in principle implies that we can build a reviewable safety case. The dilemma, however, is that we are still not able to build *sufficient* conditions for safety, because one can always argue that there is some unknown sequence of events, which would lead to an accident. With the proposed approach, we think that there are possibilities to argue that sufficient conditions can be claimed for at least some restricted parts of the production system and its controls.

The paper starts from a general discussion of threats, risk and safety, which touches on the question how safety can be built and demonstrated. We argue for the need of considering the four systems, man (people), technology, organisational and information (MTOI) all with their own modelling paradigms. We consider the design basis threat (DBT) concept, because this concept provides a set of initiating events that in many cases can give a starting point for system design. Of course it is also important to be aware of threats that are excluded from the analysis for some reason or another.

The third section considers in more detail requirements for a successful control of safety. Considering safety controls, one may separate between *state control* and *transition control*. State control represents the controls necessary to keep a system is in a safe region of a state space, and transition controls transfer the system back to a safe state if it has entered an unsafe region.

The fourth section discusses five control structures, which in the control of large complex systems are used and combined in various ways. These controls have their own characteristics, which are important to understand in modelling, designing, operating and maintaining their functionality. Controls are applied for different purposes and one may separate between main and supporting control tasks. Hierarchical controls are formed through an interconnected network of control loops that get their inputs from many diverse sources and which can influence both concrete and abstract entities.

From there we move to the general problem of modelling. To ensure a proper understanding of sociotechnical systems and their risks it is necessary to include at least four distinctly different systems, man (the M-system), technology (the T-system), organisation (the O-system) and information (the I-system). An important part of the modelling effort is to select a state space of the used models and to assess how these state spaces may be divided into three regions, one region of safe states, one region of unsafe states and one region where safety is undecided.

In the sixth section we discuss structural and mathematical prerequisites that place serious impediments on possibilities to predict system behaviour. In this section we also briefly discuss the concept of polycentric control. With this approach one may construct a set of "small worlds" for which controls can be designed and assessed. By the use of independent autonomous systems it is likely that less predictability would result, but we think that the balance nevertheless will be positive, since this approach has the benefit of building resilience into different parts of a system (Hollnagel et al., 2006).

In the seventh section we discuss management systems with the intent of bringing concepts from management science in line with the proposed control metaphor. To model the O-system in larger details, it is necessary to consider organisational structures that are defined through processes and functions. Towards the end of the section we argue, along with many others, that the feedback of experience may be the most important function within safety management. Organisational changes close the loop from feedback of experience to actual improvements.

In the eighth section, we take a closer look on implications for safety management. An important insight is that safety cannot be the only condition that influences preferences in the control loops. Effort should also be spent on how other performance criteria besides safety are prioritized and enter the controls. A specific question is to consider differences in preferences during major lifecycle phases such as design, construction, operation and decommissioning. Audits, assessments and reviews as well as regulatory oversight can be perceived as control loops that aim for obtaining indications on deviations from norms and standards to initiate correcting actions.

In spite of research efforts and development of safety management there are still a number of dilemmas that have to be addressed on a case to case basis. One is connected to limits about what we know and another is what can be considered to be safe enough. Risk profiles often have their centre of gravity at low probability, high consequence events, which lead to large uncertainties in calculated risk estimates. Selecting the focus for a modelling effort is a challenging task, but if the entire system can be covered together with the most important safety controls, at least some confidence in safety can be reached. Additional dilemmas are related to finding suitable balances in preferences as well as creating suitable models of decision making.

A conclusion of the paper is that we find the control metaphor helpful in many respects. Especially the consideration of safe and unsafe regions in a state space may provide a path forward. Polycentric control may also contribute to new insights for safety when the behaviour of an interconnected network of "safe" systems is investigated. A main conclusion is however that uncertainty in risk estimates, in spite of modelling efforts, will remain large enough to motivate an application of the precautionary principle in societal decision making.

2. Threats, risks and safety

The concepts of risk and safety are constructed through the consideration of threats to which many uncertainties are associated (Aven et al., 2011). If a threat is realised by an initiating event, it will normally come with consequences in terms of costs for the system operator (and the society). It is therefore in the interest for system operators and the society to implement and otherwise support measures, by which threats could be eliminated, isolated, controlled and/or mitigated. Risks management involve two interacting parts, an *analysis* part, where threats are identified and assessed and a *design/implementation* part, where risks are acted upon. Safety improvements may include changes in system design as well as implementing safety barriers, active safety systems and protective functions. The acceptability of building and operating potentially dangerous systems is usually controlled by society, where the operator is obliged to present a safety case with arguments for why the system can be considered safe.

2.1. Probabilities or possibilities

Quantification of risk will need assessments of the uncertainties involved. A starting point is to consider uncertainties associated to an initiating event $h \in H$ that is effecting the system at a time in-