International Journal of Industrial Ergonomics xxx (2016) 1–14

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

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

Properties for formally assessing the performance level of human-human collaborative procedures with miscommunications and erroneous human behavior☆

Dan Pan [a], Matthew L. Bolton [b, *]

[a] Department of Industrial Engineering, Tsinghua University, Beijing 100084, PR China
[b] Department of Industrial and Systems Engineering, University at Buffalo, State University of New York, Amherst, NY 14260, USA

A R T I C L E I N F O

Article history:
Received 6 June 2015
Received in revised form
22 January 2016
Accepted 4 April 2016
Available online xxx

Keywords:
Human-human collaboration
Human communication
Task analysis
Model checking
Human error

A B S T R A C T

Human-human interaction and collaboration is crucial to teamwork, where team members work together to perform tasks and share information to ensure mutual understanding. Human-human collaborative procedures are developed to ensure that relevant information is correctly heard and actions are correctly executed. Such procedures should be designed to be robust to miscommunications and other erroneous human behaviors. However, such procedures can be complex and thus fail in ways not anticipated by designers. To address this, previous efforts have used formal proof analyses to assess the robustness of collaborative procedures to miscommunications. However, these analyses only indicate strict success or failure: outcomes that fail to capture the degrees of success of collaborative procedures. Further, none of these analyses considered the interaction between miscommunications and other erroneous human behaviors. In this paper, we create specification properties to evaluate the level of success of a collaborative procedure formally. We demonstrate the use of these properties to formally evaluate realistic collaborative procedures from a nuclear power plant with and without both generated miscommunications and erroneous human behavior. We discuss the results of this evaluation and outline area of future research.

Relevance to industry: The method, performance levels, and associated specification properties allow analysts to compare the robustness of different collaborative procedures to miscommunication and attentional slips. The power of our approach is demonstrated with the nuclear power plant application. It can be easily adapted for use with collaborative procedures from other domains.

1. Introduction

With the development of industrial technology, safety-critical systems in nuclear power plants (NPP), the chemical process industry, and air transportation have become more complex. As such, their safe operation depends not only on the individual skills and knowledge or human operators, but also effective and efficient team communication and collaboration. In these sensitive systems, failures can be associated with the erroneous behavior of individual human operators (Reason, 1990) as well as human-human collaboration. For example, it is reported that in Germany, communication errors are responsible for about 10% of the workplace incidents resulting from human error (Sträter, 2003).

Of particular interest to this work is the main control room (MCR) of nuclear power plants (NPPs), where communications and collaboration among operators are essential factors for understanding how and how well MCR operators deal with abnormal or emergency situations. In particular, the performance of MCR crew under abnormal/emergency situations in NPPs is strongly affected not only by operators' cognitive processes, but also by communication and collaboration among operators. Communication error has been considered as one of the main causes of accidents and incidents in NPPs. Hirotsu et al. (2001) reported that in Japanese

---

☆ Note that *this paper is an extended version of a conference paper* (Pan and Bolton, 2015) *that was presented at HCI International* 2015. This article constitutes a significant contribution beyond this original manuscript. Specifically, it includes erroneous human behavior in the analyses; presents verification results (statespace size and verification times) for each analysis; and evaluates a new collaborative procedure (procedure 2) so that it can be compared to the original procedures (procedure 1). It also features an extended discussion.

* Corresponding author.

E-mail address: mbolton@buffalo.edu (M.L. Bolton).

http://dx.doi.org/10.1016/j.ergon.2016.04.001
0169-8141/© 2016 Elsevier B.V. All rights reserved.

Please cite this article in press as: Pan, D., Bolton, M.L., Properties for formally assessing the performance level of human-human collaborative procedures with miscommunications and erroneous human behavior, International Journal of Industrial Ergonomics (2016), http://dx.doi.org/10.1016/j.ergon.2016.04.001

NPPs, 25% of human error incidents were due to communication failure. Sträter (2003) investigated 232 operational events involving human error in German NPPs and found that roughly 10% of them which involve human errors were mainly caused by communication problems. Similar results have been observed in ground transportation (Murphy, 2001), medicine (Wilson et al., 1995), and aviation (Connell, 1996).

From these investigations and analyses, we can conclude that maintaining reliable communication and human behavior is essential to secure the safety of large, complex systems. If team members could perform and collaborate better, the safety of many systems would be improved. So far, standard human-human collaborative procedures and communication protocols have been used to ensure effective and efficient collaboration in many safety-critical systems. For example, operation crews in MCR of nuclear power plants use communication protocols to diagnose problems and execute emergency operations (Kim et al., 2010). However, there is concurrency between human operators and parts of procedures. The concurrency creates complexity and thus potentially induces unanticipated interactions between operators. Further, humans are fallible. They can perform protocols erroneously by incorrectly performing their parts of the procedure or by miscommunicating information to team member. Therefore, it can be difficult to evaluate the safety of human-human collaborative procedures using conventional analyses methods, like experimentation and simulation that can miss unexpected conditions and interactions.

Formal methods offer proof-based analysis techniques capable of considering all possible interactions. While formal methods have been used to evaluate machine communication protocols, the existing approaches (Bochmann and Sunshine, 1980; Sidhu and Leung, 1989) are ill-suited for use with human-human collaborative procedures for several reasons. First, humans behave in different ways from machines. Humans follow tasks as opposed to machine code and human-human communication must be contextualized as part of a task (Traum and Dillenbourg, 1996). Second, humans are more flexible than machines and are thus fallible in different ways. Third, human collaborative procedures are inherently less fragile than machine communication protocols because of the looser dynamics of human-human communication. As such, the outcome of human-human collaboration may represent degrees of success beyond a simple binary one (correct or incorrect). For example, if two persons are attempting to collaboratively diagnose a problem, it is problematic if they end up with only one reaching the correct conclusion. However, this is better than if both reach the same incorrect conclusion because the incorrect conclusion has a better chance of being identified and corrected as humans continue to collaborate.

Procedures for both collaborative and non-collaborative situations have been assessed formally to determine if they are safe, even with generated erroneous behavior and/or miscommunications (Bolton et al., 2013; Bolton, 2015). However, these analyses are still limited in that they have not considered the way erroneous human behaviors (e.g., attention error) and miscommunications can interact. Further, much like machine communication protocols, they only consider the binary success of human-human collaboration. This is constraining because it does not give analysts the tools they need to fully evaluate the robustness of such procedures. Therefore, an approach is needed to account for both individual erroneous behavior and miscommunication between individuals while giving analysts metrics for assessing the degrees of a procedure's success in different conditions.

In this paper, the approach in Bolton (2015) is extended to allow an analyst to model human collaborative procedures in the context of a task analytic modeling formalism and use model checking to evaluate the degrees of a procedure's success even with erroneous human behaviors. Before presenting the method, we cover the background material necessary for understanding it. We then present the method and describe how it was realized. In doing this, we use an NPP diagnosis case study to frame our analyses and illustrate how our approach can be applied to the evaluation of realistic safety-critical, human-human collaborative procedures. Finally, our results and future works are discussed.

2. Background

2.1. Formal method

Formal methods are tools and techniques for proving that a system will always perform as intended (Clarke and Wing, 1996). Model checking is an automated means of performing formal verification, checking whether a system model adheres to specifications (Clarke et al., 1999). A system model is a representation of a system's behavior in a mathematical formalism such as a finite state machine. A specification is a formal description of a desirable property about the system, usually in a temporal logic. Model checking works by exhaustively searching a model's statespace for violations of the specification. The result of this is documented in a verification report which contains either a confirmation if the model adheres to the specification or a counterexample if it does not. A counterexample lists the incremental model states that resulted in the specification being violated. This can be used by analysts to address the discovered failure.

There are a variety of temporal and modal logics that have been used to express specifications. The most common one, and the one used in the presented work, is linear temporal logic (LTL) (Emerson, 1990). LTL allows one to reason about the relationship between different states and/or variables over ordinal time and assert properties about all of the paths through a model. It does this using model variables; basic Boolean logic operators including $\wedge$, $\vee$, $\neg$, $\Rightarrow$, and $\Leftrightarrow$; and temporal operators (Table 1).

While formal methods have traditionally been used in the analysis of computer hardware and software systems, a growing body of work has been investigating how to use them to evaluate human factors issues (Bolton et al., 2013). However, when it comes to issues of human-human communication and coordination, there has been very little work. The Concur Task Trees formalism (Paternò et al., 1997) has been extended to allow for the modeling of human-human coordination and communication, where communications could have different modalities (synchronous or asynchronous, point-to-point, or broadcast), and used to formally evaluate pilot and air traffic control radio communications during runway operations using different shared task representations (Paternò et al., 1998). Although this method is useful, it did not easily distinguish between separate and shared operator tasks, nor did it account for potential miscommunications and operator perceptual or cognitive errors. Both limitations were addressed by the Enhanced Operator Function Model with Communications (EOFMC).

Table 1
Temporal operators of LTL for specification.

Name	Operator	Interpretation
Global	$G\,\Phi$	Φ will always be true.
Next	$X\,\Phi$	Φ will be true in all next states.
Future	$F\,\Phi$	Φ will eventually be true.
Until	$\Phi\,U\,\Psi$	Φ will be true until Ψ is true.

Please cite this article in press as: Pan, D., Bolton, M.L., Properties for formally assessing the performance level of human-human collaborative procedures with miscommunications and erroneous human behavior, International Journal of Industrial Ergonomics (2016), http://dx.doi.org/10.1016/j.ergon.2016.04.001