Engineering Applications of Artificial Intelligence 64 (2017) 295–301

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

A hybrid reinforced learning system to estimate resilience indicators

Simon Enjalbert *, Frédéric Vanderhaegen

Univ Lille Nord de France, France
UVHC, LAMIH, France
CNRS UMR 8201, F-59313 Valenciennes, France

ARTICLE INFO

Keywords:
Resilience engineering
Learning
Man–machine systems

ABSTRACT

This paper describes a learning system based on resilience indicators. It proposes a hybrid learning system to estimate Human–Machine System performance when facing unprecedented situations. Collected data from various criteria are compared with data estimated using the local and the global resilience indicators, to give both instantaneous and over-time Human–Machine System states. The learning system can be composed of two different, separate reinforcement functions; the first allowing reinforcement of its own system knowledge and the second allowing reinforcement of its estimation function. When used together in a hybrid approach, the resilience indicator estimation should be improved. The learning system is then applied in a simulated air transport context and the impact of each reinforcement function on resilience indicator estimation is assessed. The hypothesis on performance of hybrid reinforcement learning is confirmed and it provides better results than those obtained by the knowledge based reinforcement or the estimation based reinforcement alone.

1. Introduction

Ouedraogo et al. (2013) defined resilience as the positive ability of a Human–Machine System (HMS) to recover from or adapt to critical situations. The recovery function consists of getting back to the previous normal functioning state and the adaptation function aims to provide the system with a new stable functioning state. A large amount of research has been performed in research laboratories about system safety and security in transport or industry based on this concept (Orwin and Wardle, 2004; Pérez-España and Arreguın-Sánchez, 2001; Enjalbert et al., 2013; Cacciabue et al., 2013). Some of this research involves assessment based on various criteria. These system evaluation criteria mainly concern human or machine behaviours or their effects, or the occurrence or consequences of external perturbations. These effects or perturbations relate for instance to human workload (Vanderhaegen, 1997), to human errors (Lin et al., 2015), to the quality or the production of services (Polet et al., 2009), and to the quality of cooperation or learning activities (Vanderhaegen et al., 2006). Therefore, resilience emerges in a risk management process and relates to the system capacity to survive both planned and unexpected hazardous events (Enjalbert et al., 2011). Unprecedented situations are defined as events with a very low frequency of occurrence and/or which may have catastrophic consequences for HMS.

This paper focuses on the learning system developed to estimate resilience indicators. The reinforcement functions of the learning system concern reinforcement of the system knowledge and reinforcement of the estimation parameters. This hybrid approach has been developed and tested on a flight simulator during an in-flight refuelling activity involving a team composed of four people. Several unexpected events with potential catastrophic consequences are incorporated and data collected on HMS during unprecedented situations are used by the hybrid reinforced learning system to estimate the local and the global resilience indicators.

In the second section of the paper, the need for resilience indicators and the principles of the reinforced iterative learning approaches are presented in order to introduce the contribution of the present work. In the third section, the generic architecture of the learning system is detailed with specific focus on reinforcement functions. Finally, a validation example showing the impact of reinforcement functions on resilience indicator estimation is described and the effectiveness of hybrid reinforcement is demonstrated.

2. Learning approaches and resilience assessment

Several concepts of learning can be found in literature. For instance, learning by imitation or observation consists in copying a given

* Corresponding author at: UVHC, LAMIH, France.
E-mail address: simon.enjalbert@univ-valenciennes.fr (S. Enjalbert).

http://dx.doi.org/10.1016/j.engappai.2017.06.022
Received 27 May 2016; Received in revised form 28 April 2017; Accepted 26 June 2017

behaviour, or a sequence or a repetition of behaviours (Chella et al., 2006; Calinon et al., 2007). When facing a new situation for which no knowledge is defined, trial-and-error based learning should be applied (Rose et al., 2014). A redundant learning system is another way to engage the learning capacity of the system (Vanderhaegen and Zieba, 2014). Cooperative learning or co-learning are then useful for exchanging data between decision makers in order to understand the learning process or to share knowledge (Doisy et al., 2014). Effective techniques, characterized by efficient self-learning and adaptivity abilities, have been employed to construct learning systems (Xu and Yan, 2004; Liu et al., 2013; Norrlöf and Gunnarsson, 2005; Wiering and van Hasselt, 2008). Many of these involve reinforcement learning or reinforced learning. Reinforcement learning is usually applied for repetitive tasks, in order to minimize tracking errors. If the error reduction is successful, the reinforcement is based on a reward for managing knowledge. Other authors prefer using the vocabulary of reinforced learning because their interest is not limited to repetitive tasks and error tracking reduction. Vanderhaegen et al. (2011) focused on the learning from human errors in order to provide human operators with decision support tools.

In this study, the learning approach objective is to estimate missing or immeasurable data from Human–Machine System facing unprecedented situations. In the first section, a theoretical analysis based on extended State of the Art of Iterative Learning Control (ILC) systems is proposed. Then, in a second section, indicators based on resilience criteria for HMS are developed. Finally, these indicators are adapted to reinforced learning approaches.

2.1. Iterative learning control systems

The feedforward process aims at assessing the future possible decisions regarding the current system states and the management of the previous ones. The feedback aims at recovering possible erroneous knowledge, at refining knowledge or at creating new knowledge (Vanderhaegen, 2010). So the feedforward-feedback mechanism that consists in using the current knowledge related to previous activities in order to calculate the future ones. A great number of research works have proposed feedback and/or feedforward controllers using different methods in order to reach the mentioned objectives. There are frequency based approach (related to iteration frequency) or temporal based approach (related to timing process).

Iterative Learning Control (ILC) systems are used to benefit from the repetitive nature of the tasks as experience gained to compensate for the poor or incomplete knowledge of the plant model and disturbance. The repeatability of the task determines the learning ability of the ILC. Current (e_i) in Eq. (1) and previous (e_{i-1}) in Eq. (2) tracking errors , and previous input u_{i-1} are used to assess the current input u_i in Table 1. The recursive process of ILC technique to assess the current characteristics and to improve tracking control performance in batch processes is given in Eq. (3). The formalism can be seen as a generalization of the previous ones; the control is done regarding the previous errors at certain level because of limited memory capacity. A feedback-feedforward structure for the trajectory tracking of a linear Direct Current motor is given in Eq. (4). The same structure for sharp tracking control of a manipulator robot, by employing a saturated input γ which limits the control input within a reasonable bound, was also proposed. The corresponding learning control updating law is given by Eq. (5). The class of non-linear systems to which the proposed learning scheme can be applied is then extended. A combined feedback-feedforward controller and disturbance observer designed for a direct drive motion control was proposed in Eq. (6). The digital disturbance observer is included in the proposed feedback-feedforward control structure to compensate for disturbances (friction and cogging effects). Finally, a framework for the assessment of the consequences of human errors based on learning and prediction of the actions of a human operator is given in Eq. (7) in Table 1. These processes are modelled by using the iterative learning control concept and by integrating it in a feedforward-feedback approach.

Table 1
Different formalisms for feedforward and/or feedback based learning control.

References	Formula	
Xu et al. (2004) Ojha et al. (2017) Geng et al. (2017)	$u_i = u_{i-1} + G_{feedforward}(e_{i-1})$	(1)
Xu et al. (2004) Radac and Precup (2016)	$u_i = u_{i-1} + G_{feedback}(e_i)$	(2)
Lee and Lee (2007)	$u_i = u_{i-1} + G_1(e_{i-1}) + \cdots + G_p(e_{i-p})$	(3)
Lee et al. (2000)	$u_i = u_{i-1} + G_{ff}e_{i-1} + G_{fb}e_i$	(4)
Jang et al. (1995)	$u_i = \gamma v_i = \gamma(u_{i-1} + G_{ff}e_{i-1} + G_{fb}e_i)$	(5)
Yan and Shiu (2008)	$u_i = u_i^{ff} + u_i^{fb} - u_i^d$ $= G_{ff}(e_{i-1}, u_{i-1}) + G_{fb}e_i - G_d(e_{i-1}, u_{i-1})$	(6)
Vanderhaegen et al. (2009) Polet et al. (2012)	$u_i = e_i + G((e_{i-1}, u_{i-1}), \dots, (e_0, u_0))$	(7)

ILC has become a competitive control method through the development of different learning controllers for many applications, essentially in robotic operations, chemical processes and motor drive machines. Initially the ILC input signal is formed using the error from previous iterations, *i.e.*, the input u_i is computed using the previous input u_{i-1} and e_{i-1} in so-called Previous Cycle Learning (PCL) in Eq. (1) or recursively $e_{i-1}, \dots, e_{i-p}$ in Eq. (3). Several authors have computed the input u_i using the current tracking error e_i in so-called Current Cycle Learning (CCL) in Eq. (2). Then, it has been proposed to combine the current error, e_i with the previous one e_{i-1}, when forming u_i in Eqs. (4)–(6). This approach leads to a causal relationship between the current error and the input signal. It can be seen that PCL and CCL are functioning a complementary manner with the aim to improve the control performance through Previous and Current Cycle Learning (PCCL) structure, complementary role of feedback and feedforward structures.

The formalisms, summarized in Table 1, are used to deal with machines processes control (optimize robot or motor motion) during repetitive tasks – mostly tracking errors performance control – by managing a static knowledge. These control processes are not applied to problems involving humans and do not manage knowledge in unexpected or unprecedented situations. An extended approach by using the previous couples $((e_{i-1}, u_{i-1}), \dots, (e_0, u_0))$ was proposed with feedforward-feedback learning control systems having their updating laws mostly depending on current and/or previous errors in Eq. (7). The originality of this model is that it is applied to HMS with the aim to predict human errors. It combines feedforward-feedback processes and use predefined knowledge that is reinforced or corrected regarding the observed previous couples.

A State of the Art has been realized to compare different structures of the feedforward and/or feedback Iterative Learning Control systems in order to select the more appropriate one or to build an efficient one, for improving knowledge on known situations and for creating knowledge related to new situations. Therefore, the proposed article extends the Iterative Learning Control concept by proposing a hybrid reinforced learning structure that reinforces the learning by controlling two criteria of learning errors: errors between knowledge and error between predictions by taking into account matrices of data instead of vectors of data. Moreover, this new structure is applied to predict resilience indicators.

2.2. Resilience indicators

The proposed learning contribution should be able to estimate instantaneous and over-time HMS states, called respectively the local and the global resilience indicators. These indicators are based on several criteria such as the success level of a given task, the safety level of this task or the human workload in terms of interactions with the technical systems. For an iteration i and k criteria of resilience, the vector denoted U_{ki} in Eq. (8) is based on two indicators, the local indicator, u_{ki}, and the