

Contents lists available at ScienceDirect

Safety Science

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



Extended FRAM model based on cellular automaton to clarify complexity of socio-technical systems and improve their safety



Takayuki Hirose*, Tetsuo Sawaragi

Department of Mechanical Engineering and Science, Kyoto University, 615-8540 Kyoto, Japan

ARTICLE INFO

Keywords: Socio-technical systems Functional resonance analysis method (FRAM) Complexity Resilience engineering

ABSTRACT

The safety of socio-technical systems is intractable by traditional approaches, such as a simple cause-effect analysis. It would be no longer effective only to eliminate potential hazards or to explain why things go wrong (i.e., the limitation of the approach called Safety-I). The safety instead could be improved by validating why everyday operations go right and by verifying their resilience, based on an alternative idea of Safety-II. The objective of this research is to develop a tool enabling us to validate and verifying safeties based on the functional resonance analysis method (FRAM). FRAM would be an effective way to analyze the safety of complex socio-technical systems. However, FRAM only provides a conceptual methodology, and further extensions and implementations are needed. This work first presents an extended FRAM model based on a classical idea of cellular automaton, and the result of applying our model to a steel production management problem is provided. Steel production line reveals a quite complicated process consisting of many linked workstations where a production flow might be varied frequently. Wherein flexible delivery decisions are required, including changing the supply chains themselves. There exists empirical know-how to tackle those, but these remain as tacit knowledge, and there has been no way to validate and verify their effectiveness under specific conditions. The results of applying our extended FRAM model to this example provide several insights concerning the characteristics of experienced workers' operations to handle and manage the process's complexities, such as harnessing, phase transformation, and identifying critical points attaining resilient operations in terms of the entropy of the process status. We also discuss how these findings contribute to the safety management of the other socio-technical systems based on the findings of the analysis.

1. Introduction

Innovative technological developments and the complications of society have been making systems more and more complex, resulting in the creation of System of Systems (SoS) (Selberg and Austin, 2008). Those SoS that involve human factors, technical factors, organizational factors, and working environment are known as socio-technical systems, typical examples of which include the operations of airlines, railways, and supply chains. They play vital roles in our daily lives, and their safety is one of the most critical issues currently facing society.

We often face challenges when it comes to thinking about the safety of these systems with traditional approaches such as why-because analysis. For example, it can be difficult to determine precisely what caused an accident, because although individual factors often seems too trivial to be the cause, interactions among them could cause severe and unexpected outcomes. By the same token, it can be difficult to know why operations are going well; although operators of socio-technical

systems have some empirical knowledge or information for better operations, it remains unclear why they work well.

This difficulty is mainly due to the complexity of socio-technical systems. While agents of the systems obey local rules, they are also interrelated with each other, resulting in emerging outcomes. The interactions bring about unexpected outcomes that are difficult to explain with a logical combination of each agent's behavior. That is why it is difficult to follow how they work with the linear cause-effect relationships. The safety of socio-technical systems must be investigated on the basis of not only the linear cause-effect relationships but also the idea of emergence.

Resilience engineering (Hollnagel et al., 2006) has recently been attracting attention in this respect Resilience in this context means a system's ability to function while avoiding catastrophes even though unexpected or unwanted events degrade the safety. Resilience engineering aims at enhancing such ability, and it is realized by understanding why everyday works go right and improving it, which is an

E-mail address: hirose.takayuki.27v@kyoto-u.jp (T. Hirose).

^{*} Corresponding author.

idea of Safety-II (Hollnagel et al., 2013). However, there are no standard tactics to validate them due to the complexity of socio-technical systems, and this is making discussions of resilience engineering confusing. Therefore, it would be essential to develop a tool that enables us to visualize the safety by validating why things go right. For this purpose, the functional resonance analysis method (FRAM) could contribute to it as the first step of resilience engineering.

FRAM (Hollnagel, 2004; Hollnagel, 2012) is an effective way to investigate the safety of socio-technical systems because it provides guidelines to understand the complex dynamics emerging from the interactions among the various components that make up the systems. FRAM functions are defined as what must be done to achieve a specific goal, and FRAM investigates how the variabilities and the interactions among functions influence their safety. However, FRAM is a method rather than a model (Hollnagel, 2012), which means that it basically provides only the concept for how to determine the safety of sociotechnical systems. Further elaborations are required for its practical use.

The objective of this research is to develop a tool that enables us to grasp the visualized safety of socio-technical systems and to validate why things go right based on the idea of Safety II. To address this issue, Hirose et al. (2017) developed a FRAM simulator that can evaluate how variabilities in functions and interactions influence their safety quantitatively. They applied the simulator to an actual air crash accident and investigated how the standard operation procedure (SOP) was degraded by variabilities of the working environment, resulting in the crash. We further extended the FRAM model based on theirs in this research, using the idea of cellular automaton (Neumann et al., 1966), which is a fundamental approach to understand the characteristics of complex systems (Mitchell, 2009; Johnson, 2009). Also, we simulated the operation of steel production lines with the extended model, as an example, to validate why their operations successfully work The general steel production industries are forced to face more uncertainties than other industries. Those uncertainties make it challenging to anticipate the effectivity of operations on the production outcomes, such as attained delivery dates or expected production loads (Shioya et al., 2015). Also, although there exists some empirical knowledge or information for better operations of steel production, it is difficult to explain why the various processes work well. We, therefore, examine the effect of empirical knowledge provided by an experienced engineer working in the steel production industry and uncover several insights about the characteristics of complexity, including harnessing, phase transformation, and the relationships between critical point and entropy.

In this paper, we present our extended FRAM model and then discuss a case study through which we elaborate on various insights to improve the safety of socio-technical systems from the perspective of system complexity. We close the paper with a discussion of how these findings can contribute to the safety of socio-technical systems.

2. Safety of socio-technical systems

2.1. Four principles of safety

Ideas about safety have been changing with the development of technologies and the complications of society. Safety is traditionally assumed to be ensured by eliminating malfunctions of mechanical components or human errors. This idea originated in a time when artifacts were simpler and thought to be more independent than they are currently. However, complicated behavior in socio-technical systems has appeared due the non-linear interactions among the various agents in a system, and these are becoming important factors affecting the safety and brittleness of artifacts. Notably, the interaction of variabilities between task performance in humans or machines under a particular working environment is considered to have a significant effect on the safety of a system (Hollnagel, 2004). Hollnagel (2012) explains it with four principles that play essential roles in the safety of



Fig. 1. Traditional idea of success and failure: they stem from different sources.

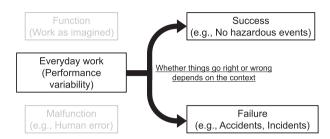


Fig. 2. Equivalence of success and failure: the same factors could lead to various outcomes depending on contexts.

socio-technical systems.

The first principle is the equivalence of success and failure. Conventionally, the state of a system is thought to be bimodal: function or malfunction, as shown in Fig. 1. Therefore, the sources of success and failure are completely independent of each other, and eliminating the sources of failure (such as malfunction of machines or human errors) is assumed to enhance the safety accordingly.

However, the situation has been changing due to the increased complexity of both mechanical systems and society. Specifically, system states can be multimodal in the sense that they are variable and flexible between "function" and "malfunction". In other words, success and failure are equivalent because, depending on specific contexts, they possibly come from the same source, as shown in Fig. 2. Therefore, it is no longer feasible to consider only success or failure individually, nor to seek just one "root cause" of accidents or success stories.

The second principle is that of approximate adjustments. It is desirable that systems be operated by means of accurately executed operational procedures that adhere to the instructions or rules issued by organizations (e.g., the company or the government). However, in reality, there may exist variabilities of the working environment caused by temporal conditions such as available resources (e.g., time) or the existence of simultaneous goals to be attained, as shown in Fig. 3. These conditions tend to create a trade-off in the operations of socio-technical systems; although operators are fundamentally required to execute predetermined procedures precisely, they should perform them in a

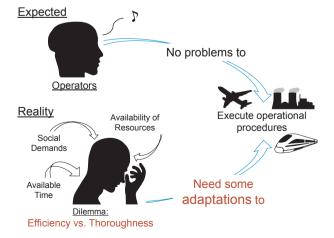


Fig. 3. Approximate adjustments: how the difference between work-as-imagined and work-as-done is caused.

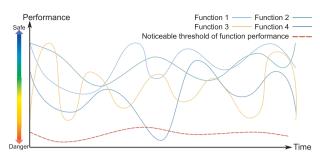


Fig. 4. Schematic illustration of functional resonance.

more flexible way to cope with the situation (e.g., the deviation from standard operation procedures (SOPs)). This dilemma is known as the efficiency-thoroughness trade-off (ETTO) principle, and in most cases, the operations of socio-technical systems cannot escape it (Hollnagel, 2004). In other words, operations are always having to adapt to a given situation. This adaptive behavior comes in the form of approximate adjustments, and it generates variabilities of task performance in humans/machines (Fig. 4).

The third principle is emergence. The operations of socio-technical systems often face unexpected outcomes, most of which are difficult to explain with decomposition and causality—a typical way to investigate what has happened. In such cases, the outcome is said to be emergent rather than resultant; if the unexpected outcomes are tractable based on known processes or developments, we regard them as "resultant." Here, note that emergence does not mean that something happens "magically"; it means that something happens whose process cannot be explained by cause-effect relationships.

The fourth principle is that of functional resonance, which is introduced to support the third principle, emergence, implying that the interactions of each component could go beyond simple cause-effect relationships. Traditionally, "resonance" refers to phenomena such a system oscillating with large amplitude when the oscillating components come together with specific frequencies. For example, random noise could make weak signals exceed the detection threshold with the principle of resonance, known as stochastic resonance, which is usually used to enhance the sensitivity of a device. We use this physical phenomenon as an analogy to clarify how the interaction between variabilities can lead to emergent phenomena. The interaction of variabilities described in the second principle could resonate in a specific context, and some variabilities that are usually too weak to notice could go beyond our noticeable threshold. This phenomenon is functional resonance, which is one of the leading causes of unexpected outcomes of socio-technical systems.

Therefore, the conventional safety analysis methods based on cause-effect relationships are no longer feasible because success and failure are equivalent. The safety of socio-technical systems depends highly on variabilities resulting from approximate adjustments, which in turn cause emergent phenomena driven by the principle of functional resonance. That is why we need new approaches supported by the above four principles to fully understand the safety of complex socio-technical systems. The two approaches with the most potential are resilience engineering (Hollnagel et al., 2006) and FRAM (Hollnagel, 2004, 2012). Our objective in the current work is to make socio-technical systems more resilient by means of FRAM.

2.2. From Safety-I to Safety-II: concept of resilience engineering

There are two ways to ensure the safety of systems: one is to eliminate potential hazards or why things go wrong, and the other is to focus on and pursue why things go right. Hollnagel et al. (2013) called these ways Safety-I and Safety-II, respectively, and has advocated a paradigm shift from Safety-I to Safety-II.

Safety-I and Safety-II are based on the ideas shown in Figs. 1 and 2,

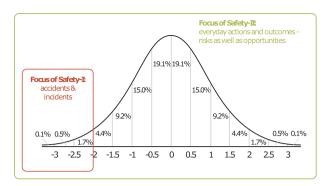


Fig. 5. Focus of Safety-I and Safety-II (Hollnagel et al., 2013).

respectively. Fig. 1 implies that the causes of success and failure can be thoroughly distinguished, and removing or weakening the causes of adverse outcomes can enhance safety. This is the essence of Safety-I, which aims to achieve a stationary state in which "there is no unacceptable risk and an undesirable state never happens." In contrast, Fig. 2 implies that success and failure could result from the same source, and we should not focus on success or failure individually. Also, as shown in Fig. 5, the frequency of unacceptable events such as accidents is perhaps very low, and almost all of the events relate to normal daily activities or are successful events. For this reason, focusing on and pursuing why things go right could be an effective way to enhance safety, which is the principle of Safety-II. Here, it is important to note that Safety-II does not neglect to consider why things go wrong; it considers all of the events as a whole, as shown in Fig. 5.

Resilience is a crucial aspect based on Safety-II to improve the safety of socio-technical systems, and it has been attracting attention especially since the 2011 Tohoku earthquake and tsunami, which triggered the Fukushima Daiichi nuclear disaster. In its original definition, resilience means (1) the ability to become recovered, happy, or prosperous again after a difficult situation or event, or (2) the ability of a substance such as rubber to return to its original shape after it has been deformed or bent (Longman Dictionary of Contemporary English). This characteristic is essential in the field of safety engineering, and resilience in the field is now regarded as "a system's ability to function while avoiding catastrophes even though disturbances such as variabilities have degraded the safety."

Resilience engineering (Hollnagel et al., 2006) seeks ways to enhance the resilience of systems. The resilience of socio-technical systems depends on how well they can adjust to a specific context surrounding them. In other words, verification and validation of approximate adjustments, which is the second principle of safety, play vital roles in designing resilient socio-technical systems. However, there are no standard tactics for achieving them, which makes any discussion of resilience engineering quite complicated. We need to pursue an understanding of why things go right in order to clarify why systems can be resilient. Resilience engineering hinges on Safety-II in that sense, and FRAM could contribute to the enhancement of resilience engineering.

2.3. Functional Resonance Analysis Method (FRAM)

The functional resonance analysis method (FRAM) (Hollnagel, 2004; Hollnagel, 2012) is a method to understand how the variabilities and their interactions among functions lead to various outcomes, including unexpected events. FRAM can be used to investigate actual events caused by non-linear interactions of variabilities. It can also be used to simulate the effect of variabilities existing in socio-technical systems by means of the "If-Then Exercise.".

FRAM starts by identifying functions. Functions in FRAM are defined as what has to be done to achieve a specific goal; each task described in a manual or procedure is a typical example. Also, the functions are supposed to have six aspects, as listed in Table 1. These aspects

Table 1
Six aspects of function.

Description
Input or trigger of functions
Outcome of functions
Conditions that must be satisfied before functions are carried out
What is consumed by functions (e.g., fuel, energy, labor force)
What supervises or restricts functions
Time required to accomplish functions

connect functions, and a socio-technical system is represented as a network of them. If the purpose of FRAM is to investigate anomalous events, functions can be obtained from databases related to those events, such as accident reports. Also, if the purpose of FRAM is to simulate the effect of variabilities with an "If-Then Exercise," various methodologies such as hierarchical task analysis (HTA) (Kirwan and Ainsworth, 1992) are available.

After the functions are identified, potential coupling arises among them. Potential coupling refers to the dependencies that can exist among functions. An easy way to find these couplings is to consider the linguistic relationships between the output of one function and five aspects (input, precondition, resource, control, time) of the rest of the functions. For example, suppose that we have three functions: TO START CAR, TO RELEASE FOOT BRAKE, and TO SHIFT FROM PARK TO DRIVE. The function TO START CAR is triggered by releasing the foot brake, and the gear must be shifted from park to drive before departure. Therefore, the input of TO START CAR can be "Foot brake is released," and the precondition of this function can be "The gear has already been shifted from park to drive." On the other hand, the output of to release foot brake and to shift from PARK TO DRIVE can be "Foot brake is released" and "The gear has already been shifted from park to drive," respectively. These outputs correspond to the input and precondition of TO START CAR, building a potential network of functions, as shown in Fig. 6. The manner of finding potential couplings is currently qualitative rather than quantitative, and a more systematic way is desired (Hollnagel, 2012).

Given a set of functions with their potential coupling, a process called instantiation is carried out. In this process, a network with a specific pattern of dependencies among functions, called an instance, can be obtained, as shown in Fig. 7. For example, a standard operating procedure such as "work-as-imagined" might be one of the instances, and other procedures that emerge as "work-as-done" are other possible instances. The instantiation process can provide various "actual works," which are the result of approximate adaptations, as instances.

FRAM is a method to investigate the effect of variabilities existing in a specific instance. Variabilities in functions of the instance propagate through the network, and they interact with variabilities in other functions and influence the performance of those functions. Moreover, they resonate in a specific context and can change the situation

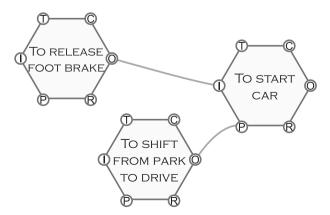


Fig. 6. Example of potential coupling: starting a car.

significantly, causing unexpected outcomes. This is the fundamental idea of FRAM and is how it currently operates. It is different from the conventional safety analysis method in the sense that FRAM examines what will happen when some variabilities come together rather than analyzing events on the basis of decomposition and causality.

However, as stated earlier, FRAM is a method rather than a model (Hollnagel, 2012), which means it can only provide the concept of how to understand the safety of socio-technical systems. Moreover, the definitions of essential entities of FRAM—such as variabilities, propagations, and their interactions—are still too ambiguous; we need to set clearer definitions and implement specific models if we want to enable practical use of FRAM. To this end, we have been developing a FRAM model by introducing numerical definitions of those entities.

There has already been extensive research devoted to the construction of FRAM models. Duan et al. (2015) integrated FRAM with a computer tool, model checking (Clarke et al., 1987), to define the interaction of variabilities among FRAM functions so that it can automatically search the potential paths that could lead to hazards. Yang et al. (2017) extended this model by adopting the Simple Promela Interpreter (SPIN) (Holzmann, 2004), which is a kind of model checking, to illustrate functional resonance or emergence. Patriarca et al. (2017) proposed a semi-quantitative FRAM model based on Monte Carlo simulation that can highlight critical FRAM functions and the critical links among them. Slater (2017) represented a network of FRAM functions as a Bayesian network to evaluate the state of each function with probabilistic values. Lee and Chung (2018) tried to quantify the effect of variabilities existing in human-system interaction and proposed a model comprising a FRAM instance and a network of operators based on the heterogeneous network theory to suggest the critical part of human-system interaction and support the management of those variabilities.

Compared with the approaches above, our FRAM model is characterized as an approach that provides qualitative comprehension of the safety of socio-technical systems on the basis of quantitative criteria. Any discussion of the safety of socio-technical systems is apt to contain some ambiguity in terms of evaluating the safety, and it would be preferable to develop a new method that enables evaluation both from the qualitative and quantitative perspectives. However, the FRAM models proposed so far tend to focus on only the qualitative approach or the quantitative, not both. Our proposed FRAM model takes the balance into account: specifically, it can support the qualitative interpretation of the safety of target systems and simultaneously quantify the degree of the danger to which each FRAM function gets anomalous.

3. Extended FRAM model based on cellular automaton

Our extended FRAM model is based on work by Hirose et al. (2017), who newly introduced numerical definitions of FRAM entities such as variabilities, their propagations, and their interactions by applying Fuzzy CREAM, which is an advanced model of the cognitive reliability and error analysis method (CREAM) (Hollnagel, 1998). They also introduced several equations to formulate the propagation and interaction of those variabilities in Fuzzy CREAM. While our proposed FRAM model adopts these ideas as well, the significant difference is that, in our model, the order of these processes is determined in accordance with the concept of cellular automaton (Neumann et al., 1966). This modification enables FRAM to reveal the dynamical transition of safety in each function.

In this section, we provide two essential explanations. First, we describe the primary mechanism of the FRAM model, namely, how the variabilities of the working environment induce a variability of function, and how they interact with each other when implementing the previous model. Second, we present the entire process of the extended FRAM model using that mechanism, which is the new part.

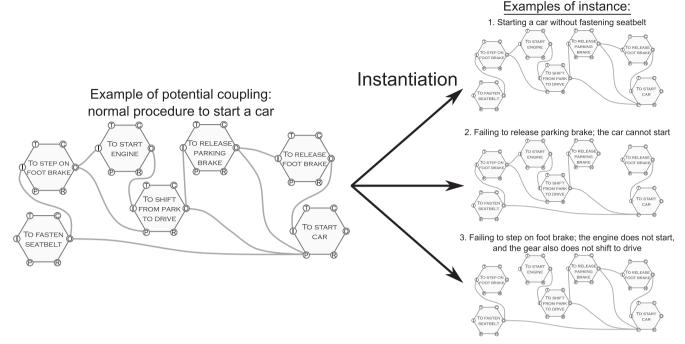


Fig. 7. Example of instantiation: normal procedure to start a car.

3.1. Primary mechanism of FRAM model

3.1.1. Numerical definitions of variabilities with fuzzy CREAM

To begin with, numerical definitions of variabilities are required, and Fuzzy CREAM is adopted in this model. An advanced model of the cognitive reliability and error analysis method (CREAM) (Hollnagel, 1998), Fuzzy CREAM enables variabilities of the working environment and functions to be represented quantitatively.

CREAM is the second-generation form of human reliability analysis, which investigates how events are going well on the basis of common performance conditions (CPCs) evaluations. Conventionally, in first-generation HRA (e.g., the technique for human error rate prediction (THERP)), human error was thought to stem from inherent deficiencies and the fact that humans naturally fail to perform tasks, just the same as machines or structures can fail. However, an extensive study revealed that the contextual conditions under which a work is performed have a significant influence on human failure, which led to the development of the second-generation HRA.

The ATHEANA methodology has been proposed as a representative method of the second-generation HRA (Barriere et al., 2000). Bearing in mind that the environment and the surrounding context may affect the behavior of a human operator, ATHEANA takes account of what is known as error-forcing contexts (EFCs), which are then combined with performance shaping factors (PSFs). It can analyze the occurrence of an actual unacceptable event, and the result can clarify how to improve safety. CREAM is an extended method of ATHEANA.

In the CREAM method, E. Hollnagel referred to contextual conditions collectively as common performance conditions (CPCs) and then defined and classified them into nine factors: "Adequacy of organization", "Working conditions", "Adequacy of man-machine interface", "Availability of procedures", "Number of simultaneous goals", "Available time", "Circadian rhythm", "Adequacy of training and experience", and "Crew collaboration quality". After that, two more CPCs, "Available resources" and "Quality of communication", were added to the original nine when FRAM was proposed for the first time (Hollnagel, 2004), implying that a set of CPCs can be modified ourselves depending on the situation. Each CPC contains various CPC levels and effects, as shown in Table 2. For example, if the CPC "Working conditions" is rated as "Advantageous", it has a "Positive" effect on the

Table 2
Examples of CPC levels and effects.

CPC	Level	Effect
Working condition	Advantageous Compatible Incompatible	Positive Not Significant Negative

situation. All CPCs are evaluated in the same way. The number of CPCs whose effect is found to be "Negative" or "Positive" is obtained in the analysis.

The effect of CPCs is updated according to dependencies among CPCs after a set of all CPCs' effects is identified. Some CPCs are interrelated with each other, and their relationships are shown in Table 3. If an effect of CPCs in the left column of Table 3 is "Not Significant", and more than three or four CPCs in the right column are "Positive" or "Negative", the effect of CPCs in the left column also becomes "Positive" or "Negative", respectively.

On the basis of the set of CPC effects and the chart shown in Fig. 8, the control mode is identified. The control mode represents how events are going and is determined as one of four degrees: Strategic, Tactical, Opportunistic, or Scrambled. Strategic represents a state in which people can carry out their tasks the most efficiently; namely, they can consider the global context and choose the next action on the basis of sophisticated strategies. In the Tactical control mode, people can carry out their tasks on the basis of planning, and can follow procedures or rules to some extent; the scope of the plan is somewhat limited, which means crucial points for the planning might be missed. Opportunistic represents a situation where people choose the next action be relying heavily on the salient features of a current context rather than more stable strategies or goals. Scrambled is a state in which people are too upset to choose a proper action; an extreme case here would be the state of momentary panic. Also, each control mode is related to an interval of probability of action failure (PAF), as shown in Table 4. Here, note that the chart in Fig. 8 carries the premise that the weights of CPCs, which represent the significance of those CPCs for a subject of CREAM analysis, are all equivalent.

CREAM has been applied to investigate human reliability in many

Table 3
Dependencies among CPCs.

Working conditions
Number of simultaneous goals
Available time
Available time
Adequacy of MMI, Availability of procedures
Working conditions, Adequacy of MMI, Availability of procedures, Number of simultaneous goals
Available time
Working conditions, Adequacy of MMI, Availability of procedures, Number of simultaneous goals, Adequacy of training and experience,
Circadian rhythm
Crew collaboration
Adequacy of organization, Circadian rhythm, quality Quality of communication

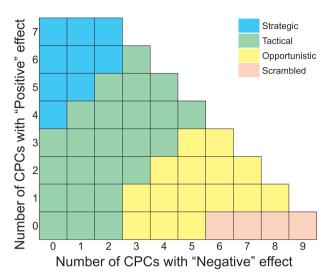


Fig. 8. Relation between CPC effect and control modes.

Table 4PAF intervals with respect to control modes.

Control Mode	Intervals of probability of action failures
Strategic	0.50×10^{-5}
Tactical	0.10×10^{-2}
Opportunistic	0.010
Scrambled	0.10

fields so far. Akyuz and Celik (2015) adopted CREAM to assess human reliability along with the cargo loading process onboard liquefied petroleum gas (LPG) tanker ships. They developed a human error assessment approach based on CREAM and applied that to the monitoring of the crew's cognitive actions or attitudes during cargo operations onboard LPG tankers. They concluded that the approach could be applied to any other critical operational processes. Zhou et al. (2017) evaluated the human reliability of seafarers performing their onboard operations. In their work, they introduced eight customized CPCs for better capturing the essential aspects of the working situations and conditions for tankers. They also built a model using the Markov method to estimate a quantified human error probability (HEP) and identified the result of the error probability intervals are limited within the tolerant ranges using the original CREAM. Zhou et al. (2017) incorporated CREAM and Monte Carlo simulation into fault tree analysis (FTA) (Lee et al., 1985) to evaluate a Liquefied Natural Gas (LNG) carrier spill accident. They constructed a modified FTA model for LNG spill accidents during LNG carriers' handling operations and introduced the CREAM model to predict human errors in the operations. Their results were synthesized in the end so that Monte Carlo simulation can provide risk as intervals of probability. Moreover, in the nuclear energy field, Yoshida et al. (2002) applied CREAM to evaluate the effectiveness of accident management (AM) which is prepared for unexpected emergencies based on probalistic probabilistic safety assessment (PSA). They developed a new method to quantify the decision-making failure probability of an

emergency organization facing an emergency of nuclear power plant, which provides an AM strategy, by using THERP and CREAM. Then they applied it to the case of a typical pressurized water reactor (PWR) plant. In conclusion, they found their method can work effectively even if the analyst is not a professional of human reliability engineering field and is applicable to other fields.

Moreover, several studies (Konstandinidou et al., 2006; Yang et al., 2013) have introduced fuzzy logic theory into the original method to make CREAM continuous and quantitative; these are generally called Fuzzy CREAM. In Fuzzy CREAM, membership functions of CPC levels whose support set is CPC score are defined. CPC score is a continuous value varying from 0 to 100 that represents the status of the CPC: the higher the CPC score, the better the CPC status. These membership functions also represent the degree of matching between a specific CPC score and a particular CPC level, varying from 0 to 1.00. Four membership functions of the control modes, whose support set is the logarithm of PAF, are defined in the same way. Then, fuzzy linguistic rules, which are IF-THEN rules between combinations of CPC levels and a specific control mode, are defined, e.g.,

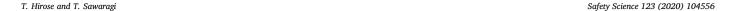
IF
$$S_1 = Compatible \ AND \ S_2 = Efficient \ AND \ \cdots \ AND \ S_m = \cdots \ THEN \ C = Strategic,$$

where S_i denotes the level of the i-th CPC, m is the total number of CPCs, and C represents the control mode $(1 \le i \le m)$. With the above items, a conclusion fuzzy set of the control mode is obtained by calculating how the antecedent matches the consequent in those If-Then rules.

There exist several models based on this fundamental idea. For example, in the former study (Konstandinidou et al., 2006), 46,656 fuzzy rules are constructed by hand using the chart in Fig. 8, and a conclusion fuzzy set is obtained by the min-max inference technique (George et al., 1995). Also, in the latter study (Yang et al., 2013), the relative weight of CPCs is defined, and the Bayesian network of CPCs calculates the belief degrees of each control mode.

In the proposed FRAM model, the *weighted CREAM model* (Ung, 2015), which is also one of the Fuzzy CREAM models, is adopted, for two reasons. First, our method considers the concept of CPC weight; while the weight is different case by case, it is regarded as equal in many cases for the sake of simplicity. Second, the table in Fig. 8 is not necessary, as the chart in Fig. 8 would not be available if the CPC weights were not equivalent. A model that satisfies both of these requirements is quite rare, and the *weighted CREAM model* is one of them. The algorithm consists of the following four steps:

Step 1: Definition of membership functions for linguistic values of CPC levels. The first step of this model is to define membership functions. Examples are shown in Fig. 9. (a) and (b) show examples of the membership functions of CPC levels. (c) shows the membership functions of control modes, along with the logarithm of the probability, whose base is 10, which is used for their support set because the lower limit of PAF is assumed to be 0.50×10^{-5} according to the original CREAM (Table 4). Ideally, the membership functions should be designed with statistical data and/or with the knowledge of experts. However, in this paper they are all regarded as the simple triangular functions shown in Fig. 9 for the sake of simplicity.



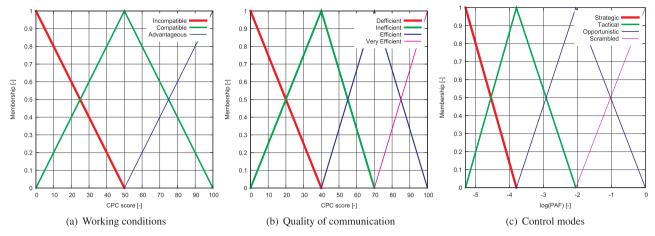


Fig. 9. Examples of membership functions.

Step 2: Construction of fuzzy rules

Fuzzy rules are constructed with a systematized process in this step. Ideally, a rule should be obtained with the statistical data and/or the knowledge of experts. However, since each of the nine CPCs has three or four CPC levels, as shown in the example in Fig. 9, tens of thousands of combinations of CPC levels are obtained as antecedents, i.e., the IF-part of the rule. For example, two levels are identified with respect to a CPC score in the case of Fig. 9, generating 29 combinations of CPC levels as a result of evaluating all nine CPC scores. Therefore, systematic ways to distribute those combinations to a specific control mode, THEN-part, and obtain an IF-THEN rule as shown below are required.

$$IF \begin{bmatrix} S_{1,1} = \cdots & AND & \cdots & AND & S_{1,m} = \cdots \\ S_{2,1} = \cdots & AND & \cdots & AND & S_{2,m} = \cdots \\ \vdots & & \ddots & & \vdots \\ S_{l,1} = \cdots & AND & \cdots & AND & S_{l,m} = \cdots \\ \vdots & & \ddots & & \vdots \\ S_{n,1} = \cdots & AND & \cdots & AND & S_{n,m} = \cdots \end{bmatrix} THEN \ C = C_k$$

Here, n is the total number of combinations of CPC levels belonging to the k-th control mode: C_k $(1 \le k \le 4)$. Also, $S_{l,i}$ represents a level of the i-th CPC in the l-th combination of CPC levels in the IF-part $(1 \le l \le n)$.

An index I^l is introduced (0 $\leq I^l \leq$ 100) to distribute a combination of CPC levels to a specific control mode. The index is defined as

$$I^{l} = \sum_{i=1}^{m} A_i^{l} \cdot w_i, \tag{1}$$

where A_i^l is the significance of the i-th CPC level in the l-th combination of the IF-part ($0 \le A_i^l \le 100$). It is defined as a value on the abscissa where the membership function reaches 1.00. For example, the significance of "Advantageous", "Compatible", and "Incompatible" is 0, 50, and 100, respectively in Fig. 9 (a). Also, w_l is the normalized relative weight of the i-th CPC ($0 \le w_l \le 1.00$) with the following equation:

$$w_i = \frac{W_i}{\sum_{i=1}^m W_i},\tag{2}$$

where W_i is the relative weight of the *i*-th CPC set by analysts ($W_i \ge 0$). I^l is regarded as a percentage on the abscissa in Fig. 9 (c); the value on a specific point of the abscissa, SV^l , is identified by following equation ($-5.30 \le SV^l \le 0$).

$$SV^l = -5.30 \times \frac{I^l}{100}.$$
 (3)

A combination of CPC levels, i.e., the IF-part, belongs to a specific control mode C_k , i.e., the THEN-part, depending on the intervals listed in Table 5 to which the SV^l belongs. Here, the intervals are obtained by

Table 5 Relationships between intervals of log(PAF) and control mode.

Control Mode	Strategic	Tactical	Opportunistic	Scrambled
Interval of log(PAF)	[-5.30, -3.80]	(-3.80, -2.90]	(-2.90, -1.03]	(-1.03, 0]

applying the OR operation of fuzzy theory for Fig. 9(c).

Step 3: Acquisition of Fuzzy Conclusion

In this step, a concluding fuzzy set of the control mode is obtained by the calculation of μ^{C_k} , the degree of matching for each control mode, which is obtained by

$$\mu_l^{C_k} = \sum_{i=1}^m \mu_{l,i}^{C_k}(x) \cdot w_i, \tag{4}$$

$$\mu^{C_k} = \frac{\sum_{l=1}^n \mu_l^{C_k}}{n},\tag{5}$$

where $\mu_{l,i}^{C_k}(x)$ is the value of the membership function corresponding to the level of the *i*-th CPC in the *l*-th THEN-part whose linguistic consequent is C_k , all of which vary from 0 to 1.00. Also, x is the CPC score ranging from 0 to 100.

The concluding fuzzy set $\mu(y)$ is obtained by using μ^{C_k} , which is defined as

$$\mu(y) = \min(\max(\nu^{C_1}(y), \mu^{C_1}), \max(\nu^{C_2}(y), \mu^{C_2}), \dots, \max(\nu^{C_k}(y), \mu^{C_k}), \dots),$$
(6)

where $\nu^{C_k}(y)$ is the membership function of the k-th control mode shown in Fig. 9(c) (0 $\leq \nu^{C_k}(y) \leq 1.00$), and y equals log(PAF) varying from -5.30 to 0..

Step 4: Defuzzification

The concluding fuzzy set is transformed into a crisp value by the following defuzzification process:

$$CV = \frac{\int_{D} y \cdot \mu(y) \, dy}{\int_{D} \mu(y) \, dy},\tag{7}$$

where CV is the crisp value of log(PAF) and D is the domain of integration. In other words, the crisp value corresponds to the center of gravity of the fuzzy set.

3.1.2. Implementing mechanism of FRAM function

The above Fuzzy CREAM model is applied to each function of FRAM to create the mechanism of each; the variabilities are defined as the parameters of Fuzzy CREAM, and equations that formulate their

propagation and interaction are newly introduced. In other words, FRAM is implemented as a network of the Fuzzy CREAM models driving in each function.

The numerical definition of variabilities is based on the concept of second-generation HRA, in which the state of a working environment or context influences the performance of tasks. Moreover, the contributing factors of the working environment are classified into the eleven elements as CPCs in CREAM. Thus, the variability of the working environment can be modeled as a dynamic transition of CPC scores; the variability of functions can also be regarded as their consequence: change of a "continuous" control mode or a crisp value of PAF. Therefore, each function is assumed to have a vector of CPCs in this model, and the variability of the working environment and function are modeled as the transition of CPC scores and PAFs of the functions, respectively.

The definitions enable us to formulate dependencies among functions or how a variability in an upstream function propagates to its downstream functions as follows:

$$x_{i,down}^{t+1} = \frac{PAF_{up}^{t}}{PAF_{up}^{t+1}} \times x_{i,down}^{t}, \tag{8}$$

where $x_{i,down}^{t+1}$ and $x_{i,down}^t$ are the updated scores and original scores of the i-th CPC in the downstream function, respectively. Also, PAF_{up}^t and PAF_{up}^{t+1} referring to the PAF value of a particular upstream function before and after the PAF value, respectively, have been changed by the Fuzzy CREAM process. Eq. 8 represents that some specific CPC scores in the downstream functions decrease if the PAF in an upstream function increases, and vice versa.

The definition of variabilities also makes it possible to formulate dependencies among the CPCs in Table 3, as follows:

$$x_i^{t^*+1} = x_i^{t^*} + \sum_j (x_j^{t^*} - x_i^{t^*}) \times w_j,$$
 (9)

where $x_i^{t^*+1}$ is an updated score of the *i*-th CPC in the left column of Table 3, and $x_i^{t^*}$ is its original score. Also, $x_j^{t^*}$ is the score of the *j*-th CPCs listed in the right column of Table 3, and w_j is the normalized weight of the CPC.

The unitary mechanism of one function is implemented as shown in Fig. 10 on the basis of the above definitions and equations, and those units are connected as shown in Fig. 11. That is, manual interventions change the original CPC scores as a trigger or variabilities coming from upstream functions first. The scores are then updated according to the dependency among CPCs by Eq. 9, which is the input of the Fuzzy CREAM process. The change of CPC scores induces the change of PAF in the function as a result of the Fuzzy CREAM process. The variability propagates to downstream functions on the basis of Eq. 8, changing the CPC scores of downstream functions. Thus, the resultant scores of CPCs being affected by variabilities coming from upstream functions are

determined both by the aspects of a downstream function and by characteristics of their upstream functions.

3.2. FRAM simulation based on cellular automaton

Now that we have shown the unitary mechanism of one function, we are ready to explain a new process in which those units interact with each other. The process is implemented with reference to the idea of cellular automaton, which enables us to visualize the dynamical change of function variabilities.

Cellular automaton was initially introduced by (Neumann et al., 1966). In trying to represent data processing in a machine that mimicked the self-reproducing that goes on in life, he discovered a kind of simulation carried out on a grid. In this process, the state of each cell on the grid is supposed to change every moment depending on the states of surrounding neighborhoods. Moreover, simple local rules to update each cell interact with each other and keep on changing the state of those cells. This is known as automaton on the grid in parallel automatically. What is interesting with this simulation is that the interaction of simple local rules could bring about a very complex macroscopic behavior: cellular automaton can execute basic logical operations AND; OR; NOT and work as if it is a computer. However, the original model that John von Neumann came up with was too complicated for practical use since each of the cells is supposed to have 29 states on the grid. More simplified models have since been built.

Conway's Game of Life (Berlekamp et al., 1982) is one of the best-known models of cellular automaton. In this model, every cell is supposed to have only two states, dead or alive, and they change their state simultaneously in parallel according to the following simple local rules (Fig. 12).

<u>Rule 1</u>: An alive cell dies in the next generation if it is surrounded by less than two alive neighbor cells, as if it were caused by underpopulation.

<u>Rule 2</u>: An alive cell can survive in the next generation if it is surrounded by two or three alive neighbor cells.

<u>Rule 3</u>: An alive cell dies in the next generation if it is surrounded by more than three alive neighbor cells, as if it were caused by overpopulation.

<u>Rule 4</u>: A dead cell can revive if it is surrounded by exactly three alive neighbor cells, as if it were a reproduction.

The local transition of each cell creates unusual global beshavior; a number of characteristic patterns can be observed in the process, including complex behavior, as if it is a computer.

These simulations are now applied to various fields to investigate complex phenomena. One of the most famous examples is the simulation of traffic flow to investigate how dynamic traffic behavior, such as traffic jams, will go on (Fukui and Ishibashi, 1996; Qian et al., 2017; Ruan et al., 2017). In other cases, they are also applied to simulations of

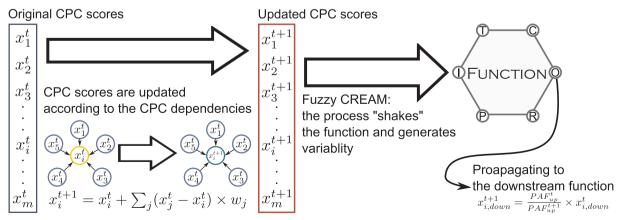


Fig. 10. Unit mechanism of one function.

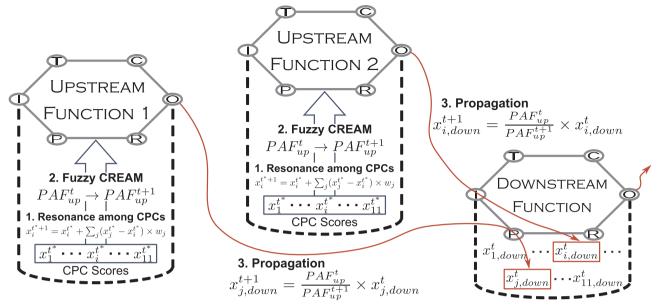


Fig. 11. Connections of units shown in Fig. 10.

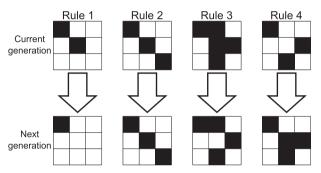


Fig. 12. Rules of Conway's Game of Life.

pedestrian group behavior during emergency evacuation (Fu et al., 2015; Lu et al., 2017). Moreover, the cellular automaton can also be applied to the field of materials science to investigate the dynamic behavior of materials such as deformation (Liu et al., 2017). The simulation of a cellular automaton is indeed applicable to a wide variety of fields, and it is also the case with the safety of socio-technical systems.

We introduce this mechanism of cellular automaton into our extended FRAM model. In this model, FRAM functions are regarded as cells, and the calculation processes shown in Fig. 10 correspond to the simple local rules of cellular automaton. Moreover, every function in a FRAM network runs and repeats those processes simultaneously in parallel; thus, a dynamic changing pattern of the PAFs of the functions is obtained. This is the significant difference compared to the previous model, which had no consideration of time transitions or parallel processing of functions. The extended FRAM model is expected to represent complex behaviors similar to cellular automaton.

The entire flow of this model is shown in Fig. 13. The process starts with setting the initial parameters, including functions, their aspects, their dependencies, instantiation, and so on; the details will be shown in the next section with an actual example. Then, the process moves on to the manual operations, where CPC scores in functions or dependency among functions can be modified. The manual operation of setting the initial values of CPS scores for the functions triggers the main process during which states of functions change on the basis of the mechanism shown in Fig. 11, similar to Conway's Game of Life. In other words, three main calculations—CPC dependency, Fuzzy CREAM, and propagation of variabilities among functions—are carried out in this order in

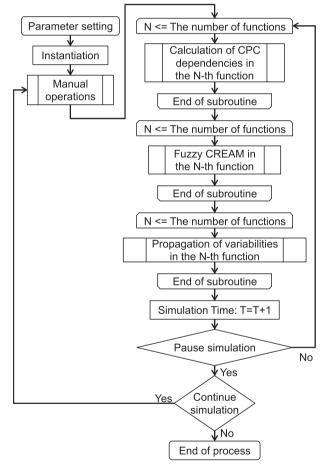


Fig. 13. Flow chart of extended FRAM model.

one generation. This is iterated until the calculation process is cut in, letting the process go back to the manual operations again. Note that the process continues automatically after the manual operations.

4. Case study: simulation of steel production line with extended FRAM model

4.1. Target of simulation: steel production line

Steel production typically uses a kind of socio-technical system, and the operations involved are often too complex to anticipate. The steel industry is generally said to contain far more uncertainties than other industries, and it is difficult to predict operations such as delivery period or production load (Shioya et al., 2015). This is mainly due to the number of product variants with which the steel industry must contend; their production processes highly depend on the specifications of those products as well as on their quality. In addition, the processes are affected by quantities of in-process inventory, system troubles, and maintenance, all of which can be related to human operators, machines, organization, and the working environment. These issues combine to make the process of steel production too complex to be anticipated.

To build a simulation model for this process, we first investigated empirical knowledge. We found that proper adjustment of the rate of direct delivery can improve material flows of a supply chain in actual steel production, according to an engineer working at a steel production company. Here, the rate of direct delivery is the ratio of what is sent directly to a downstream production process to what is sent to a storage space in the production line. However, no one can systematically explain why the adjustment is valid; how the adjustment works is so complex that no one can follow the process. The primary purpose of this simulation is to investigate the mechanism and determine the features of such emergent outcomes.

4.2. Initial setting of simulation

The simulation starts with the initial setting shown in Fig. 13, which needs to be done manually. It includes settings for both functions and CPCs. We also need to come up with a simulation scenario. The details are as follows.

4.2.1. Functions and their dependency

Fig. 14(a) shows the target system of this simulation, which is a typical production process of steel plates. The starting point of the process is importing raw materials. These materials go through production processes such as steel making, continuous casting, and rolling. While they are transported among the processes, some materials, inprocess inventories, and products are temporarily sent to storage spaces.

The processes in Fig. 14(a) are converted into a more abstracted configuration to determine the FRAM functions, and Fig. 14(b) shows the physical dependency of each abstracted process. This provides five FRAM functions first of all: TO IMPORT RAW MATERIALS, TO PREPARE CARS, TO TRANSPORT, TO PROCESS RAW MATERIALS/IN-PROCESS INVENTORIES, and TO SHIP. We also consider two new functions: TO SEND TO STORAGE SPACE and TO PREPARE STORAGE SPACE, which are related to the storage space and the dotted lines in Fig. 14(b). Therefore, a total of seven functions were defined in this simulation, with their details listed in Table 6. In addition, their potential coupling was defined as shown in Fig. 15, which is used as an instance in this simulation. It should be noted that some of the function names in Table 6 and Fig. 15 are slightly different from those in Fig. 14(b) for convenience.

4.2.2. CPC Belonging to Each Function

All functions shown in Table 6 and Fig. 15 are assumed to have a set of CPCs. In this simulation, a new set of CPCs is introduced for more specialized simulation of steel production. Specifically, the four original CPCs—Training and experience, Man-Machine interaction, Circadian rhythm, and Organization factor—were replaced with "Quality of materials," "Adequacy of lot size," "Timeliness," and "Adequacy of direct delivery rate." The four new CPCs are closely related to the operation of

steel production and have a significant effect on it, according to an engineer working for a steel production company, while the four original CPCs seems trivial or much less significant. Besides, too many CPCs can make the simulation too complicated, leading to the improper setting of initial parameters or to results that do not make sense. In total, eleven new CPCs were introduced in this simulation, which are shown in Table 7 along with their weights. Index numbers 1–7 in the second row correspond to that of the functions in Table 6.

The CPC weights were defined using a paired comparison process that is a part of the analytic hierarchy process (AHP) (Saaty, 1990). This is because the CPC weight is based on qualitative relationships between a FRAM function and its related CPCs, and it would involve too much subjectivity if it were defined intuitively. AHP was originally a methodology to support decision making (e.g., buying a car) on the basis of multiple criteria (e.g., price, fuel efficiency, size). In the process, one paired comparison of those criteria is carried out. The relative importance among each pair is evaluated with integer grades, and they are regarded as the ratio of one criterion weight to the other. Those ratios yield the following matrix A:

$$\mathbf{A} = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_m} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_m}{w_1} & \frac{w_m}{w_2} & \cdots & \frac{w_m}{w_m} \end{bmatrix}$$
(10)

where w_i ($1 \le i \le m$) is the weight of the *i*-th criterion and m is the number of criteria. The following vector is a set of weight to extract from the above matrix:

$$\mathbf{w} = [w_1, w_2, \dots, w_m]^T. \tag{11}$$

The multiplication of the matrix in Eq. 10 and the vector in Eq. 11 yields

$$\mathbf{A}\mathbf{w} = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_m} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_m}{w_1} & \frac{w_m}{w_2} & \cdots & \frac{w_m}{w_m} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}$$

$$= m [w_1, w_2, \cdots, w_m]^T$$

$$\Leftrightarrow \mathbf{A}\mathbf{w} = m\mathbf{w}. \tag{12}$$

Eq. 12 suggests that the set of CPC weights can be obtained by solving the eigenvalue problem, and the set of weights corresponds to an eigenvector whose eigenvalue is the closest to *m*—the number of criteria. In this simulation, the criteria correspond to the weight of CPCs; the above pair-wise comparison process was used to obtain the sets of CPC weight in Table 7 for each function.

In addition to the weight of CPCs, dependency among CPCs needs to be defined. The dependency among original CPCs has already been provided in Table 3. However, the new set of CPCs shown in Table 7 was introduced in this simulation, which is why the dependency among them needs to be redefined. This is shown in Table 8, where CPCs in the rows are affected by the CPC in the columns; the index numbers in the first column correspond to those in the second row. Here, 0 means there is no dependency among two corresponding CPCs, and 1 means the CPC in the column has an effect on the CPC in the row.

4.2.3. Simulation scenario

A simulation scenario needs to be set and converted into a manual change of parameters, such as CPC scores or dependency among functions. The scenario was set as shown below in this simulation.

<u>Scenario</u>: There was an excess arrival of raw materials, and the flow of materials, in-process inventories, and products grew beyond the capacity of the steel production processes. As a means for overcoming the adversity, the rate of direct delivery was adjusted at a specific timing.

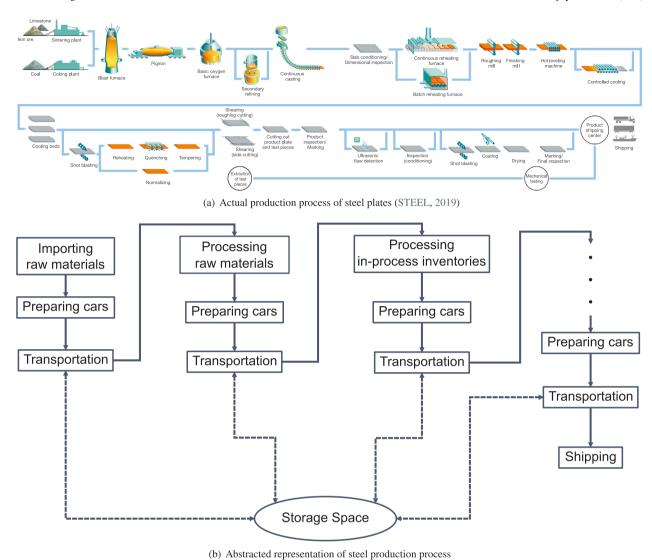


Fig. 14. Typical production process of steel plates. (See above-mentioned references for further information.)

This is converted into manual changes of FRAM entities, referred to as *Variability* and *Countermeasure*:

<u>Variability</u>: The score of CPC "Available resource" of the function to import raw materials is set to 0, which corresponds to the worst state of the CPC, at simulation time T=0.

<u>Countermeasure</u>: The simulation process is paused once at a specific timing, and the score of CPC "Adequacy of direct delivery rate" of the function TO TRANSPORT AMONG PROCESSES is set to 100, which corresponds to the best state of the CPC.

In this simulation, every timing to take a *Countermeasure* is collectively examined to determine the difference of each case. One hundred transition patterns of log(PAF) for each function is obtained as a result.

4.3. Simulation results

Fig. 16 shows three characteristic transition patterns of the log(PAF) of each function. The horizontal and vertical axes represent the simulation time and the log(PAF) of each function, respectively. The log(PAF) is equivalent to the degree of danger/instability of each function. The dotted line in the figure represents the timing when Countermeasure was adopted. For example, in the case of Fig. 16(b), log(PAF) and control mode of to transport among processes was initially -4.80 and Strategic, respectively at T=0. Then, they had been degraded to -0.68 and Scrambled by Variability until T=4.

Countermeasure was taken at this time, and the log(PAF) and control mode recovered to -4.80 and Strategic, respectively again. What is interesting with these results is that they were explicitly classified into only the three patterns shown in Fig. 16, even though 100 transition patterns of log(PAF) were obtained.

It should be noted here again that the simulation results shown in Fig. 16 are dynamical transitions of the safeties of the functions, whose plotted values are defuzzified log(PAF)s according to Eq. 7. Since these values were defined initially as fuzzy variables, it would be plausible to reinterpret the dynamical safety status qualitatively (i.e., in fuzzy linguistic values) at the end. Fig. 17 represents the transitions of the control modes of each function qualitatively with color gradations. The relationships between the values of log(PAF) shown in Fig. 16 and the values of the control mode are determined according to the definitions of the fuzzy values shown in Table 5. Hereafter, the simulation results are mainly analyzed on the basis of what Fig. 17 represents, and Fig. 16 is also referred to if necessary.

4.3.1. Interpretation of simulation results

The pattern in Fig. 17(a) shows the transitions of the control mode in each function when *Countermeasure* was adopted earlier than simulation time T=4. In this case, the control modes of each function, except to import raw materials, started their transitions towards undesirable states after *Variability* at simulation time T=0; they were

Table 6 Functions to produce steel plates.

1. To im	port raw materials	2. To tra	nsport among processes				
Input	Not Applicable (N/A)	Input	Raw materials are ready In-process inventories are processed				
Output	Raw materials are ready	Output	Transportation is ongoing/done				
Precondition	N/A	Precondition	Cars are ready				
Resource	N/A	Resource	Cars				
Control	N/A	Control	N/A				
Time	N/A	Time	N/A				
3	. To Process	4. To	prepare storage space				
Input	Transportation is done	Input	In-process inventories are processed				
Output	In-process inventories are processed	Output	Storage space is prepared				
Precondition	N/A	Precondition	N/A				
Resource	In-process inventories	Resource	Stored in-process inventories/products				
Control	N/A	Control	Flow rate of in-process inventories/products				
Time	N/A	Time	N/A				
5. T	o prepare cars	6. To send to storage space					
Input	Raw materials are ready	Input	Transportation is ongoing				
	In-process inventories are processed Storage space is prepared						
Output	Cars are ready	Output	In-process inventories/ products are sent to storage space				
Precondition	N/A	Precondition	Storage space is prepared				
Resource	N/A	Resource	Storage space				
Control	Flow rate of in-process	Control	Flow rate of in-process				
	inventories/products		inventories/products				
Time	N/A	Time	N/A				
	7. To Ship						
Input	Transportation is done						
Output	N/A						
Precondition	All processes have						
	successfully completed						
Resource	N/A						
Control	N/A						
Time	N/A						

expected to be calmed down by *Countermeasure*. However, they did not go back to the original state even after *Countermeasure* was adopted, and as such, retained an unsafe status. Moreover, the control modes of to transport among processes, to prepare storage space, and to prepare cars revealed at the end the most dangerous state of the four control modes: Scrambled. This designates a situation in which *Countermeasure* failed in preventing the occurrence of the growing danger/instability of the functions, resulting in the most dangerous/unstable state of the three patterns.

The pattern in Fig. 17(b) shows a case when *Countermeasure* was taken at simulation time T=4 or T=5. In this case, all of them went back to the safe state, Strategic or Tactical, autonomously after *Countermeasure* was adopted, in contrast to the previous case. It is remarkable here that the impact of *Countermeasure* recovered the control mode of the rest of the functions indirectly, even though *Countermeasure* itself was originally intended to improve only the control mode of to transport among functions. This pattern implies a situation where *Countermeasure* can work very effectively, and the system can go back to the safe state.

The pattern in Fig. 17(c) shows a case when Countermeasure is

adopted later than simulation time T = 5. This case is characteristic in the sense that the control mode of all functions except to IMPORT RAW MATERIALS and TO PREPARE CARS started revealing periodic transitions, even oscillations, after Countermeasure, and some of them were interrelated with each other. For example, the recovery of TO PROCESS led the control mode of to transport among processes to a dangerous/unstable state, and vice versa; a similar trend can be seen in the relationship between TO PREPARE STORAGE SPACE and TO SEND STORAGE SPACE, as well. Moreover, the control mode of to PREPARE CARS was degraded again and remained at Opportunistic, closer to Scrambled, after Countermeasure was adopted. meaning that it was difficult to prepare cars and that the resources for transportation were strictly limited. These transition patterns reveal a situation in which the recovery of one function resulted in increasing its outputs (e.g., raw materials, in-process inventories, and products) while adversely resulted in too much provision for the other functions due to the lack of cars. In other words, there were trade-off relationships among the pairs of functions because the resource for transportation was strictly limited, thus causing the oscillation patterns.

4.3.2. Factors of difference among three patterns

The factor behind the difference between Pattern 1 (Fig. 17(a)) and Pattern 2 (Fig. 17(b)) can be found in the control mode of TO TRANSPORT AMONG PROCESSES before and after Countermeasure was adopted. In the case of Pattern 2, Countermeasure succeeded in recovering the control mode of to transport among processes directly, and its effect indirectly recovered the control mode of the rest of the functions as well. On the other hand, Countermeasure failed to recover the control mode of TO TRANSPORT AMONG PROCESSES in Pattern 1, as if the "power" of the control mode to become worse was so strong that Countermeasure could not calm it down, resulting in the most undesirable state of the three patterns. The "power" in this context stemmed from the dependency among CPCs formulated as Eq. 9, since Countermeasure, or manual change of the CPC score, did not bring about any changes of the control mode in Fig. 17(a), implying that the effect of Countermeasure vanished. These two patterns suggest that the functions cannot recover without the success of Countermeasure, and that it becomes effective after simulation time T=4, when the control mode of to transport among PROCESSES reaches the most unstable state, Scrambled, and stops getting worse. In other words, the difference between Pattern 1 and Pattern 2 depends on whether Countermeasure can recover the control mode of TO TRANSPORT AMONG PROCESSES or not, and it is useless to adopt Counter*measure* earlier than the above turning point.

Similarly, the factor behind the difference between Pattern 2 (Fig. 17(b)) and Pattern 3 (Fig. 17(c)) can be found in the recovery process of the control mode of TO PROCESS. The recovery speed of the control mode after *Countermeasure* in the pattern of Fig. 17(c) was slower than that of Fig. 17(b), and it prevented the other functions from recovering. This is mainly because the log(PAF) of TO PROCESS shown in Fig. 16(c) was gradually increasing until just before *Countermeasure* was adopted; the increase was so small that it could not be observed as a change of the control mode shown in Fig. 17(c). In the end, minor changes in log(PAF) or the control mode of TO PROCESS caused the major difference between Pattern 2 and Pattern 3.

4.3.3. Summary of simulation results

The results ultimately suggest that the effect of *Countermeasure* becomes the most significant when the control mode of its target function, to transport among processes, gets as tense or unstable as the effect of *Variability* can cause. In other words, it is most effective to wait until the state of the target function becomes the most tensed or unstable state, according to Pattern 1 and Pattern 2. On the other hand, the safety of to process is also degraded little by little while waiting, and it could cause the unstable outcomes shown in Fig. 17(c) as well. Therefore, we conclude that there is a critical timing when *Countermeasure* can become the most efficient, as shown in Fig. 17(b), and *Countermeasure* at any other timing can cause chaotic outcomes, as shown in Fig. 17(a)

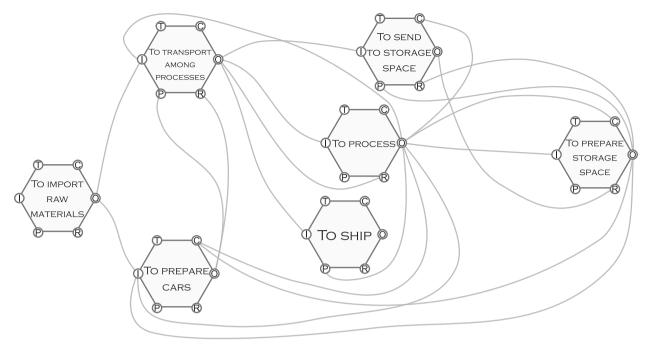


Fig. 15. Dependency among functions: instance of steel production.

Table 7
CPC weight in each function.

CPCs				Functions			
	1	2	3	4	5	6	7
1. Available resource	0.0588235	0.216650	0.152542	0.245277	0.223987	0.188505	0.223987
2. Quality of materials	0.176471	0.0190805	0.152542	0.0225458	0.0190004	0.0165634	0.0815229
3. Quality of communication	0.0588235	0.0190805	0.0508475	0.0225458	0.0815229	0.0406462	0.0248875
4. Adequacy of lot size	0.176471	0.0469760	0.0508475	0.0980026	0.0815229	0.0966109	0.0815229
5. Availability of procedures	0.0196078	0.0190805	0.0169492	0.0225458	0.0190004	0.0165634	0.0190004
6. Working condition	0.176471	0.0469760	0.0508475	0.0980026	0.0248875	0.188505	0.0815229
7. Number of simultaneous goals	0.0196078	0.0469760	0.0508475	0.0272530	0.0815229	0.0406462	0.0815229
8. Available time	0.0588235	0.216650	0.152542	0.0980026	0.0815229	0.0406462	0.0815229
9. Timeliness	0.176471	0.216650	0.152542	0.245277	0.0815229	0.142161	0.223987
10. Crew collaboration	0.0196078	0.0469760	0.152542	0.0225458	0.223987	0.0406462	0.0815229
1. Adequacy of direct delivery rate	0.0588235	0.104904	0.0169492	0.0980026	0.0815229	0.188505	0.0190004

Table 8
Dependency among the new set of CPCs.

		CPCs which have effect on									
CPCs which are affected by	1	2	3	4	5	6	7	8	9	10	11
Available resource	0	1	1	1	1	1	1	1	1	1	1
2. Quality of materials	1	0	0	0	0	0	1	1	1	1	0
3. Quality of communication	1	0	0	0	0	1	1	1	1	1	0
4. Adequacy of lot size	1	1	1	0	0	0	0	1	1	1	1
5. Availability of procedures	1	0	1	0	0	0	0	0	0	1	0
6. Working condition	1	0	1	0	0	0	0	1	1	0	0
7. Number of simultaneous	1	1	0	1	1	1	0	0	0	1	1
goals											
8. Available time	1	1	0	1	1	1	1	0	1	1	1
9. Timeliness	1	1	1	1	1	1	1	1	0	1	0
10. Crew collaboration	1	0	1	1	1	1	0	1	0	0	0
11. Adequacy of direct delivery rate	1	1	1	1	0	0	0	0	1	1	0

and (c).

5. Discussion: using complexity for the safety of socio-technical systems with the proposed model and its future prospects

There are two remarkable findings in the obtained result. One is that, even though *Countermeasure* itself was only intended to directly improve the safety of to transport among processes, its impact indirectly improved the safety of the rest of the functions as well. The other is that *Countermeasure* became effective just after its target function, to transport among processes, had gotten as tense or unstable as the effect of *Variability* could make it. There has been some discussion regarding how to predict the future evolution of complex systems and manage them accordingly, and these characteristics are somewhat related to topics that are controversial in those discussions.

5.1. First finding: efficient control of complex systems

While it might be ideal to have every artifact system, including socio-technical systems, under control, it is almost impossible to do so since socio-technical systems involve many factors in their operation, requiring infinite precision. We can find some clues to overcome this problem in the following. Israeli and Goldenfeld (2004) have suggested

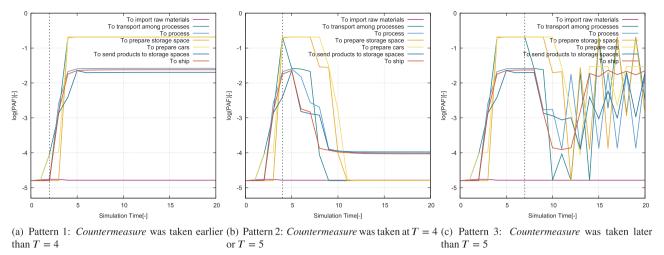
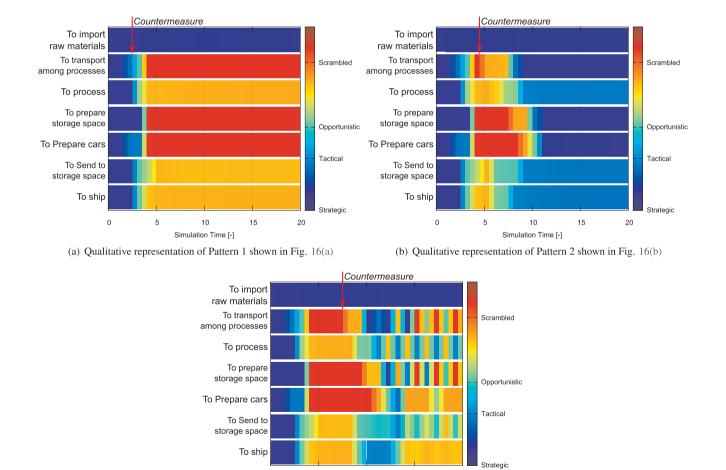


Fig. 16. Simulation results: Transition patterns of log(PAF) in each function.

that computationally irreducible physical processes can be computationally reducible at other coarse-grained resolution levels. In other words, it is possible to predict the behavior of complex systems without accounting for all of their small-scale details. Also, Smith and Johnson (2004) contended that it is not necessary to know all about target systems; only minimal knowledge and intervention on certain parts of the systems is enough to manage their future evolution. These studies

convincingly argue that it is not required to understand everything about a target complex system nor to intervene on a large part of the systems in order to control it. Instead, it is sufficient to only intervene on a small, specific part of the systems —one whose effect interacts with the others parts according to its potential dynamics and leads to desired outcomes. Moreover, complex systems can be harnessed by taking advantage of their characteristics of complexity and emergence.



 $\mbox{(c) Qualitative representation of Pattern 3 shown in Fig. 16(c)} \label{eq:pattern3}$ **Fig. 17.** Qualitative representation of simulation results: transition of control mode in each function.

10

Simulation Time [-]

15

20

5

0

The transition patterns of the control modes of the functions shown in Fig. 17(b) precisely imply the above assertions. There is no need to intervene in the CPC scores of multiple functions to overcome the adversity caused by Variability. Instead, it would be enough to adjust the CPC score of just one particular function: "Adequacy of direct delivery rate" in to transport among processes. In other words, the impact of a small intervention would propagate all around the FRAM network shown in Fig. 15, interacting with the rest of network and bringing about the most efficient and desirable outcome compared with the other two cases shown in Fig. 17. While experienced engineers in the steel production industry do know this fact unconsciously, at the same time, they also know that it is not always effective, as shown in Fig. 17. Thus, to help them overcome this confusion, tools that enable such complex, possible behaviors to be envisioned are required. This is precisely the goal of the work, and our proposed model does contribute to deliberating on the means for the indirect control (i.e., harnessing) and to clarifying why things go right. It brings to life the concept of Safety-II, thus enhancing the resilience of socio-technical systems.

5.2. Second finding: drastic change of complex systems

Complex systems can change their behavior drastically under certain circumstances (Mitchell, 2009; Johnson, 2009), which is generally called phase transition or bifurcation. The particular circumstances in which a phase transition occurs are called critical points, where small changes in specific parameters can cause qualitative changes in the behavior of a macroscopic system (e.g., between water and vapor at boiling point). One of the most famous examples of this phenomenon can be seen in a flock of birds: the entire flock can change its velocity and flying direction drastically, making it difficult for predators to attack. However, this does not mean that all the birds in the flock always fly in the same direction at the same speed; the parameters of each bird usually fluctuate and correlate with each other, causing information flow through the flock network. Bialek et al. (2014) investigated this mechanism by building a model based on the maximum entropy method, which succeeded in fully representing data observed in the real world. Their research implies that a system with high entropy could be close to a critical point, and that it can exist between order and chaos-known as the edge of chaos.

Fig. 16 reveals a similar phenomenon to the above phase transition. When the transition from Pattern 1 to Pattern 2 in Fig. 16 is taking place just after the control mode of to transport among processes, the target function of *Countermeasure* reaches its peak of log(PAF), where the instability of this system is assumed to be high. Also, such an unstable state contributed to the transition from Pattern 2 to Pattern 3: the transition is caused by the minimal increase in the log(PAF) of to process around simulation time T=5. That is, just small changes in specific parameters could cause drastic changes of system behaviors, thus demonstrating critical points or the edge of chaos. Similar behavior can be observed in the transition patterns of the control mode in Fig. 17. Ultimately, the drastic changes might also lead to undesirable outcomes, and well-considered management is essential to deal with such a situation

Although it might sound conflicting with our intuitions: adversities or unwanted outcomes must be avoided as soon as possible. There indeed exist these kinds of gaps between theory and practice in discussions about safety, and they are frontiers to explore for the future safety of socio-technical systems. What is essential to manage the safety of socio-technical systems is to consider, understand, and harness such complex dynamics of systems. The proposed FRAM model can provide us with insights about new safety methodologies, i.e., resilience engineering based on Safety-II.

5.3. Limitations and future prospects of the proposed model

There are three major points to be improved on this FRAM model.

They are closely related to both reliability and efficiency of the simulation and must be addressed in future works. On the other hand, it is also sure that those improvements bring about great impacts on this FRAM model, making it much more attractive.

The first point is that the current FRAM model requires many parameters for its initial setting of the simulation, which might consume too much energy, time, and involve subjectivities. It is especially case with the setting of CPCs because their weight needs to be defined for all functions, and the process is now depending on analysts. Also, a set of CPCs can be customized as shown in this simulation, which is also pointed by Konstandinidou et al. (2006) and Zhou et al. (2017). That is why we need to seek for some ways to collectively and efficiently define factors of working environment as CPCs, depending on subjects of the simulation. Besides, the weight of those CPCs needs to be evaluated, based on objective and automatic solutions. One possibility to solve this problem is to apply machine learning techniques to the above process in which the model is fed by real data and analyze it to create a set of CPCs and their weight. If this process is automated, it can improve both the efficiency and reliability of this FRAM model.

The second point is that the simulation results currently just provide abstract or conceptual insights about the safety of socio-technical systems. In other words, those results just provide qualitative state of the target systems, and it is difficult to know quantitative information about the safety. To address the issue, we are now assuming that another less abstracted simulations such as physical simulations are required, if necessary, to investigate what will actually happen in the real world. Moreover, the simulation of FRAM running at a higher abstraction level and some physical simulations running at a lower abstraction level need to be connected in some ways. One solution for the purpose is extracting dynamical transitions of CPC scores changing behind log(PAF) in each FRAM function. This is because they can involve more detailed information about a working environment and provide semi-quantitative criteria about the result of simulation. In other words, they can work as if they were media connecting between the higher and lower abstraction levels and bring about input and output values for both of the simulations.

The last point is that it requires some experience to interpret the results since they are still too conceptual. In this simulation, the original results of Fig. 16 were converted into qualitative color maps to provide a better understanding. However, we still need to seek other ways to represent those results with more simplified symbols so that the systems' status can be grasped at a glance. This will be important, primarily when the simulation is used in real operations in letting the operators know the status of systems immediately. For this purpose, we are currently expecting that "force dynamics" (Talmy, 1988) can be useful for visualizing such dynamics Force Dynamics is a semantic category to describe how entities interact concerning force, including the exertion of force, resistance to such a force, and the overcoming of such a resistance. According to this, the following three patterns are distinguished as different dynamical patterns:

Pattern 1: The growing effect of variability propagating through the system is currently so strong that the effect of countermeasure is hidden, and the system cannot recover from the chaotic status.

Pattern 2: The growing effect of variability propagating through the system became weaker and weaker, and this makes countermeasure effective; thus the system can overcome the variability turning out to be stable.

Pattern 3: The effect of countermeasure propagating through the system can contribute to recovering only a part of the system, while it, in turn, makes the rest of the system still in danger.

In this way, force dynamics is a useful scheme for the qualitative envisioning of the overall system status and is expecting to transfer the qualitative representations shown in Fig. 17 onto more abstract summarization contributing to the more intuitive understanding of the results.

6. Conclusion

The safety of socio-technical systems is based on four principles: equivalence of success and failure, approximate adjustments, emergence, and functional resonance. FRAM is an effective method to understand such safety and can be utilized for the improvement of these systems as resilient systems. However, FRAM is a method rather than a model, which means it provides only the concept. Therefore, FRAM needs to be extended and implemented to make it practical.

In response to this need, we developed an extension of FRAM that uses a model based on the concept of cellular automaton. The FRAM functions were built with Fuzzy CREAM and connected so that they interact with each other as in cellular automaton. The extended FRAM model enables us to visualize the dynamics of functions.

The extended FRAM model was applied to the simulation of a steel production line. In the simulation, we investigated the effect of empirical knowledge of engineers in the steel production industry by adjusting the rate of direct delivery. The results suggested that some characteristics of complex systems—namely, harnessing, phase transformation, and the relationship between critical point and entropy—played a significant role in the dynamics. This is just an individual case at this time, and the result needs to be generalized as future work. In conclusion, we wish to emphasize that it is important to consider, understand, and harness such complex dynamics of systems in order to manage the safety, and the FRAM model has the potential to make this possible.

Acknowledgement

This work is supported by Dr. Hiroshi Narazaki, an engineer of Kobe Steel, Ltd. The authors are grateful to him for his assistance for sharing a lot of insightful suggestions.

References

- Akyuz, E., Celik, M., 2015. Application of CREAM human reliability model to cargo loading process of LPG tankers. J. Loss Prev. Process Ind. 34, 39–48.
- Barriere, M., Bley, D., Cooper, S., Forester, J., Kolaczkowski, A., Luckas, W., Parry, G., Ramey-Smith, A., Thompson, C., Whitehead, D., et al., 2000. Technical basis and implementation guidelines for a technique for human event analysis (ATHEANA). NUREG-1624, Rev 1.
- Berlekamp, E.R., Conway, J.H., Guy, R.K., 1982. Winning Ways for your Mathematical Plays, vol. 2 AK Peters/CRC Press.
- Bialek, W., Cavagna, A., Giardina, I., Mora, T., Pohl, O., Silvestri, E., Viale, M., Walczak, A.M., 2014. Social interactions dominate speed control in poising natural flocks near criticality. Proc. Nat. Acad. Sci. 111, 7212–7217.
- Clarke, E.M., Grumberg, O., Browne, M.C., 1987. Reasoning about Networks with Many Identical Finite-State Processes. Technical Report. CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE.
- Duan, G., Tian, J., Wu, J., 2015. Extended FRAM by integrating with model checking to effectively explore hazard evolution. Math. Prob. Eng. 2015.
- Fu, Z., Zhou, X., Zhu, K., Chen, Y., Zhuang, Y., Hu, Y., Yang, L., Chen, C., Li, J., 2015. A floor field cellular automaton for crowd evacuation considering different walking abilities. Physica A 420, 294–303.
- Fukui, M., Ishibashi, Y., 1996. Traffic flow in 1D cellular automaton model including cars moving with high speed. J. Phys. Soc. Jpn. 65, 1868–1870.
- George, J.K., Bo, Y., et al., 1995. Fuzzy sets and fuzzy logic: theory and applications. PHI New Delhi 443–455.
- Hirose, T., Sawaragi, T., Horiguchi, Y., Nakanishi, H., 2017. Safety analysis for resilient complex socio-technical systems with an extended functional resonance analysis method. Int. J. Astronaut. Aeronaut. Eng. 2.
- Hollnagel, E., 1998. Cognitive reliability and error analysis method (CREAM). Elsevier.

- Hollnagel, E., 2004. Barriers and Accident Prevention. Ashgate Publishing Ltd.
- Hollnagel, E., 2012. FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio-technical Systems. Ashgate Publishing Ltd.
- Hollnagel, E., Leonhardt, J., Licu, T., Shorrock, S., 2013. From Safety-I to Safety-II: A White Paper. Eurocontrol.
- Hollnagel, E., Woods, D.D., Leveson, N., 2006. Resilience Engineering: Concepts and Precepts. CRC Press.
- Holzmann, G.J., 2004. The SPIN Model Checker: Primer and Reference Manual, vol. 1003 Addison-Wesley Reading.
- Israeli, N., Goldenfeld, N., 2004. Computational irreducibility and the predictability of complex physical systems. Phys. Rev. Lett. 92 074105.
- Johnson, N., $\frac{2009}{}$. Simply Complexity: A Clear Guide to Complexity Theory. Oneworld Publications.
- Kirwan, B., Ainsworth, L.K., 1992. A Guide to Task Analysis: The Task Analysis Working Group. CRC Press.
- Konstandinidou, M., Nivolianitou, Z., Kiranoudis, C., Markatos, N., 2006. A fuzzy modeling application of CREAM methodology for human reliability analysis. Reliab. Eng. Syst. Saf. 91, 706–716.
- Lee, J., Chung, H., 2018. A new methodology for accident analysis with human and system interaction based on FRAM: case studies in maritime domain. Saf. Sci. 109, 57–66.
- Lee, W.S., Grosh, D.L., Tillman, F.A., Lie, C.H., 1985. Fault tree analysis, methods and application a review. IEEE Trans. Reliab. 34, 194–203.
- Liu, Y.X., Lin, Y., Zhou, Y., 2017. 2D cellular automaton simulation of hot deformation behavior in a Ni-based superalloy under varying thermal-mechanical conditions. Mater. Sci. Eng.: A 691, 88–99.
- Lu, L., Chan, C.Y., Wang, J., Wang, W., 2017. A study of pedestrian group behaviors in crowd evacuation based on an extended floor field cellular automaton model. Transp. Res. Part C: Emerg. Technol. 81, 317–329.
- Mitchell, M., 2009. Complexity: A Guided Tour. Oxford University Press.
- Neumann, J., Burks, A.W., et al., 1966. Theory of Self-reproducing Automata, vol. 1102024 University of Illinois Press Urbana.
- Patriarca, R., Di Gravio, G., Costantino, F., 2017. A Monte Carlo evolution of the Functional Resonance Analysis Method (FRAM) to assess performance variability in complex systems. Saf. Sci. 91, 49–60.
- Qian, Y., Zeng, J., Wang, N., Zhang, J., Wang, B., 2017. A traffic flow model considering influence of car-following and its echo characteristics. Nonlinear Dyn. 89, 1099–1109.
- Ruan, X., Zhou, J., Tu, H., Jin, Z., Shi, X., 2017. An improved cellular automaton with axis information for microscopic traffic simulation. Transp. Res. Part C: Emerg. Technol. 78, 63–77.
- Saaty, T.L., 1990. How to make a decision: the analytic hierarchy process. Eur. J. Oper. Res. 48, 9–26.
- Selberg, S.A., Austin, M.A., 2008. Toward an evolutionary system of systems architecture. In: INCOSE International Symposium, Wiley Online Library. pp. 1065–1078.
- Shioya, M., Mori, J., Ito, K., Mizutani, Y., Torikai, K., 2015. Atsuita Seizou Kouki no Kakuritsu Moderu to Seizou Hyoujun Kouki Sanshutsu Gijutsu no Kaihatsu (Development of Stochastic Model of Production Time and Estimation Technology for Standard Production Time in a Plate Mill). Tetsu to Hagane (Iron and Steel) 101, 574-583. https://doi.org/10.2355/tetsutohagane.TETSU-2015-024.
- Slater, D., 2017. Modelling, Monitoring, Manipulating and Managing? Modelling Process Flow in Complex Systems.
- Smith, D., Johnson, N.F., 2004. Evolution Management in a Complex Adaptive System: Engineering the Future. arXiv preprint cond-mat/0409036.
- STEEL, N., 2019. Steel Plates. http://www.nssmc.com/product/catalog_download/pdf/A001en.pdf. Retrieved on Oct. 18th 2019.
- Talmy, L., 1988. Force dynamics in language and cognition. Cognit. Sci. 12, 49–100.
 Ung, S.T., 2015. A weighted CREAM model for maritime human reliability analysis. Saf. Sci. 72, 144–152.
- Yang, Q., Tian, J., Zhao, T., 2017. Safety is an emergent property: illustrating functional resonance in Air Traffic Management with formal verification. Saf. Sci. 93, 162–177.
- Yang, Z., Bonsall, S., Wall, A., Wang, J., Usman, M., 2013. A modified CREAM to human reliability quantification in marine engineering. Ocean Eng. 58, 293–303.
- Yoshida, Y., Ootani, M., Fjita, Y., 2002. Ninchiteki Bunnsekimoderu wo Mochiita Accident Management no Ishikettei Kagoritsu no Teiryouka (Quantification of a Decision-making Failure Probability of the Accident Management Using Cognitive Analysis Model). INSS J. 9, 183–204.
- Zhou, Q., Wong, Y.D., Xu, H., Van Thai, V., Loh, H.S., Yuen, K.F., 2017. An enhanced CREAM with stakeholder-graded protocols for tanker shipping safety application. Saf. Sci. 95, 140–147.
- Zhou, T., Wu, C., Zhang, J., Zhang, D., 2017. Incorporating CREAM and MCS into fault tree analysis of LNG carrier spill accidents. Saf. Sci. 96, 183–191.