Proceedings of the Seventh International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 15, Paper 24.2
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.15.24.2

Applicability of Generative AI to Railway Signalling Control Based on the Evolution of Fail-Safe Design

T. Itagaki[1], T. Takata[1], T. Fukuda[2], H. Mochizuki[2], S. Takahashi[2] and H. Nakamura[2]

[1]Kyosan Electric Manufacturing Co., Ltd., Yokohama, Japan
[2]College of Science and Technology, Nihon University, Japan

Abstract

Safety design in railway signalling systems has evolved along two major paths: the incorporation of safety-related non-functional requirements into systems from the outset and the prevention of the recurrence of past accidents. Historically, many such accidents have been attributed to human error. Japanese railway safety systems have achieved a high level of sophistication by reducing human involvement while continuously integrating technological advancements. As technology has advanced, the concept of safety design has also evolved significantly. Recently, attention has been directed toward constructing more rational systems based on the concept of intrinsic control, in which safety is ensured through consistency verification at system interfaces. Furthermore, recent technological developments have prompted interest in the application of Generative AI to railway signalling control systems. However, due to challenges such as limited explainability and unpredictability, its adoption remains cautious. This paper proposes an alternative approach in which Generative AI is not used for control decisions but for evaluating the consistency of data sequences, thereby enhancing safety without compromising system reliability.

Keywords: Railway signalling, fail-safe, interlocking, block system, fail-safe logic, generative AI, intrinsic control

1 Introduction

At the first stage of railway transportation system the stationmaster was required to confirm the correct position of point machines before allowing a train to enter the station. This human checking process later evolved into mechanical checking mechanisms and eventually into modern interlocking systems.

Dispatch procedures also required the stationmaster to exchange signals with the next station via telegraph to confirm that no trains were approaching. This practice established the foundation of today's block systems.

Thus, even at the dawn of Japanese railways, interlocking and block control formed the basis of safe railway operations, although both relied heavily on human attention. The inherent limitations of human based safety later led to the development of mechanical interlocking and block systems. Today, communication technology and advanced system architecture continue to push safety systems toward higher performance and reliability.

This paper shows how the concept of fail-safe design has changed throughout these and outlines our new proposal that is applicability of Generative AI for signalling control.

2 Transition of Fail-Safe Designing Technology

To apply generative AI to safety-critical signalling control systems, we will review the evolution of safety technologies in railway signalling.

2.1 Transition of Fail-Safe Technology and the Emergence of Fail-Safe Relay Circuits

There is also a fault in the signal system that instructs the driver to proceed. How can we ensure safety against the failure of signalling equipment? Various fail-safe technologies have been developed to address this issue. Figure 1 shows a mechanical signal device, known as the "ball signal". The ball signal allows the train to enter the station when the ball is in a high position. Even if the rope breaks, the ball falls in a low position to prevent the train from entering the station and ensure fail-safe operation. Conversely, if the ball in the low position rises, it would be unsafe. However, this cannot occur because it cannot rise against gravity.

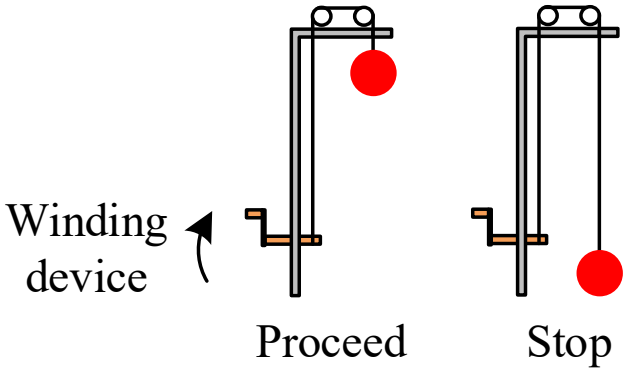


Figure 1: Example of a fail-safe device (ball signal)

In the early stages of rail transport, human errors that compromised the safety of signalling systems accounted for the majority of railway accidents. This problem was solved by the invention of a train detection sensor called a track circuit. When a human requests a route, it is processed by a relay sequence based on inputs from track circuits. Track circuits play an important role in preventing human error. Figure 2 shows a part of the driving direction circuit of the single-track line, and if the driving direction is set by the operation of the station manager of both stations, the train route can only set in that direction. In this relay sequence, whether to use the make-contact or the break-contact of the relay with asymmetric-error characteristics was important in ensuring the fail-safe nature of the system. And the know-how was passed on empirically among engineers. These safety technologies have supported the safety of signal equipment

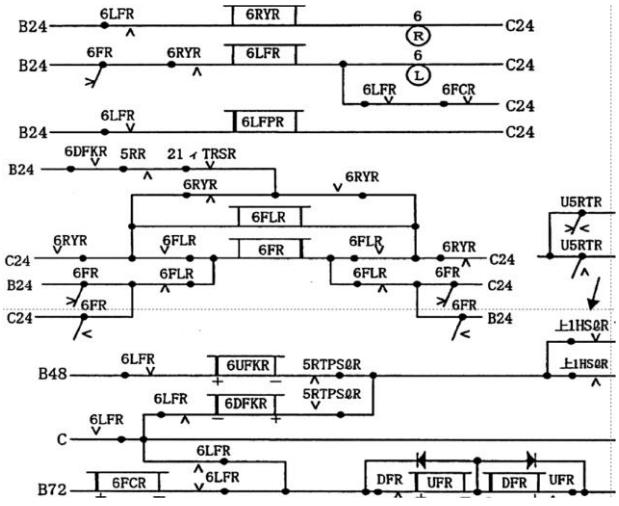


Figure 2: Part of a directional control circuit in a single-track automatic block system

2.2 Elucidation of Fail-Safe Logic Systems

Fail-safe design methods for relay circuits in signalling systems, as shown in Figure 2, have been empirically passed down, but their principles were clarified only in the mid-1960s. The origin of this research was the development of electronic interlocking devices underway at the time. The research began with the elucidation of the fail-safe properties of relay interlocking devices, one of the safety devices used in signalling systems. The principles were clarified by Watanabe, Takahashi, and others in 1965 as a method for constructing fail-safe logic systems [1]. The construction of fail-safe logic systems is based on the principle that, if a circuit is unate, a fail-safe logic circuit can be realized without redundancy using asymmetric-error devices. Furthermore, the unate property corresponds to monotonicity; that is, when a circuit is expressed in canonical sum-of-products form, each input variable appears exclusively as either a positive or negative literal, as shown in Figure 3.

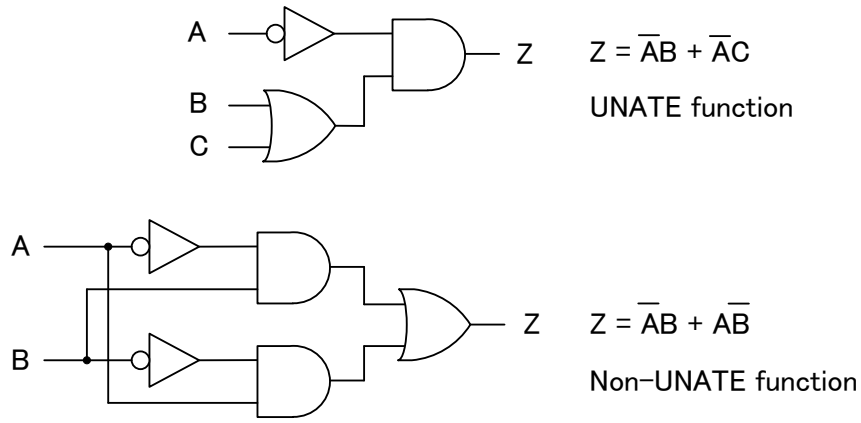


Figure 3: Examples of unate and non-unate functions

As soon as the construction method of fail-safe logic circuits was clarified, many studies on fail-safe systems were conducted at the academic level in Japan, as outlined below:

1. Methods for making non-unate circuits fail-safe
2. Methods for constructing fail-safe circuits using symmetric-error logic elements
3. Development of asymmetric-error logic elements
4. Methods for fault diagnosis of fail-safe circuits

In fail-safe signalling devices using relays, as shown in Figure 2, the selection of make or break contacts was made based on empirical fail-safe considerations. In other words, such devices at that time were designed so that each circuit element, from input to output, exhibited fail-safe properties. As illustrated in Figure 4, the entire circuit maintained fail-safe characteristics, analogous to a Baumkuchen, which exhibits the same pattern regardless of where it is cut. At the time, Okumura developed a fail-safe calculator by duplicating fail-safe logic elements [2].

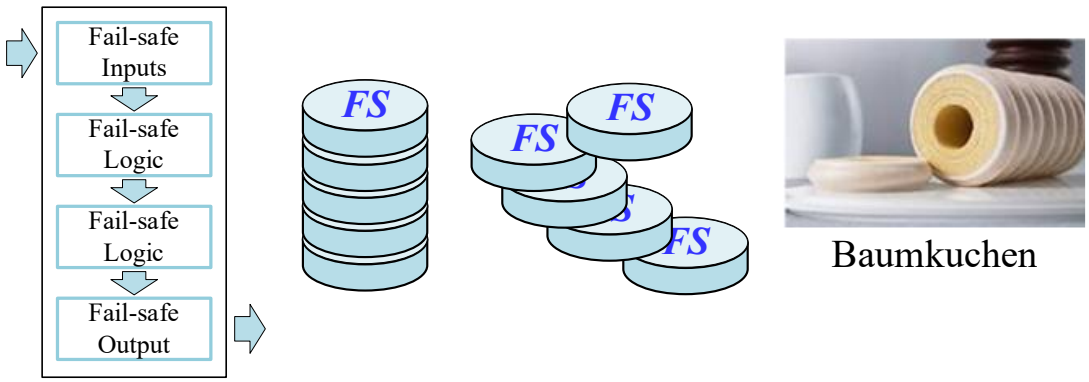


Figure 4: Structure of a relay signalling device, where each block exhibits fail-safe properties

Methods for implementing fail-safe logic in non-unate circuits and those using symmetric-error logic elements were first explored using two-wire logic, in which 1 is represented by the code word (1, 0) and 0 by (0, 1) [3]. In two-wire logic, (1, 1) and (0, 0) are invalid code words. By assigning these to physical quantities, an error output Φ can be defined, leading to the development of a three-valued fail-safe logic system [4]. In addition, various approaches have been explored, such as fail-safe coding schemes in which the number of ones is fixed at K in an N-bit code. Significant results have also been achieved in fault diagnosis, and majority voting circuits using asymmetric-error logic elements [4] are employed in computerized interlocking systems.

2.3 Utilization of Fault-Tolerance Technology

While research on fail-safe logic systems was progressing in Japan, fault-tolerant computing was being studied in the United States, particularly in the aerospace and aviation industries, where recovery from failures was difficult. The objective was to complete missions despite failures.

A representative circuit technique in fault-tolerant computing is the self-checking circuit (SCC) [5]. While fail-safe systems are typically based on binary logic (0 and 1), SCC operates in a coded space, where valid and invalid code words are explicitly defined. Within SCC, the concept of totally self-checking (TSC) was defined. Assuming a single fault, TSC ensures that a non-code word is eventually produced in the event of failure, while all preceding code words remain correct. Regardless of how to construct a TSC circuit, if the circuit is TSC, correct operation can be ensured by adding a check circuit to the output to check whether it is a non-coded word or not. However, even if the circuit is TSC, if a fault occurs and another failure occurs before the non-code word is output, the legitimacy of the subsequent processing cannot be guaranteed. On the other hand, even if multiple faults occur, the output code words remain valid until a non-code word is detected at the interface. To address this property, strongly fault-secure (SFS), the highest class of SCC, was proposed, and a construction method was presented by Dr. Nanya et al. [6]. According to this method,

the bus-synchronous fail-safe computer developed for railway systems can be regarded as conforming to the concept of SFS. Furthermore, the results obtained from fail-safe design are consistent with those derived from research in fault-tolerant computing. In addition, various approaches have been proposed, including methods for comparing processing results at program checkpoints [7] and methods based on software redundancy [8]. These approaches have been successfully implemented as highly reliable systems. Figure 5 illustrates these relationships.

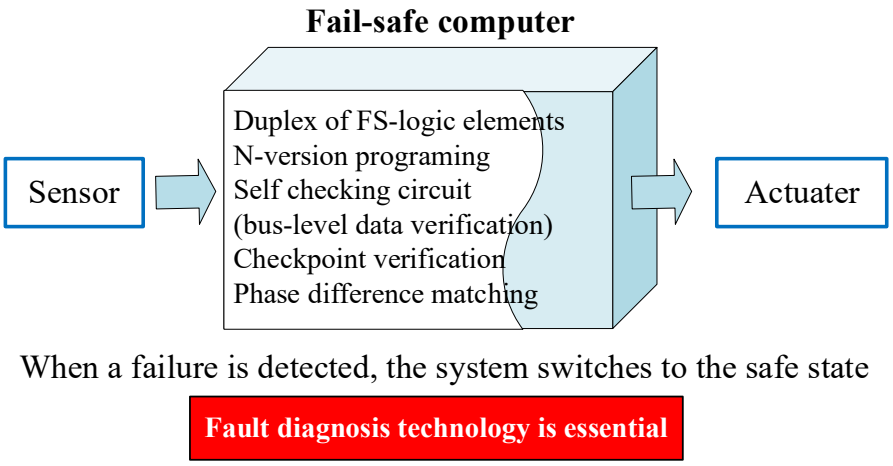


Figure 5: Examples of fail-safe computer systems with checking functions

Any of the diagnostic methods shown in Figure 5 can be used to detect discrepancies. However, it should be noted that even if a fail-safe computer is constructed using these methods, its fail-safe performance ultimately depends on hardware reliability. With regard to software, safety must be ensured through appropriate design and verification in signalling safety systems. In practice, this is achieved through thorough testing and validation of software functions. For this purpose, the application of functional safety standards [9] (Part 3) is considered effective.

2.4 General-purpose product use and systems approach

When a safety system is composed of fail-safe logic elements, fail-safe requirements can be satisfied by the circuit itself that implements the function, and function and safety are integrated. With the development of fail-safe computers and the increasing implementation of functions in software, methodologies for ensuring software safety became essential.

In signalling system software, a single-threaded scheduler is employed to execute all tasks sequentially, thereby avoiding task switching (preemption), which can introduce failures. It also requires the application of safety measures, such as asymmetric acceptance tests, during information exchange between tasks. In addition to these safety-conscious techniques, safety management throughout the software lifecycle maintains today's high-quality software.

Furthermore, when incorporating general-purpose devices into the system, it is essential to establish a rational approach to safety assurance and to build consensus among stakeholders. This concept is still under consideration; however, as shown in Figure 6, it is considered that consistency checks of data exchanged at the interfaces between components will serve as the basis. In addition, several methodologies based on current technologies are being investigated, and the results will be reported in a separate publication

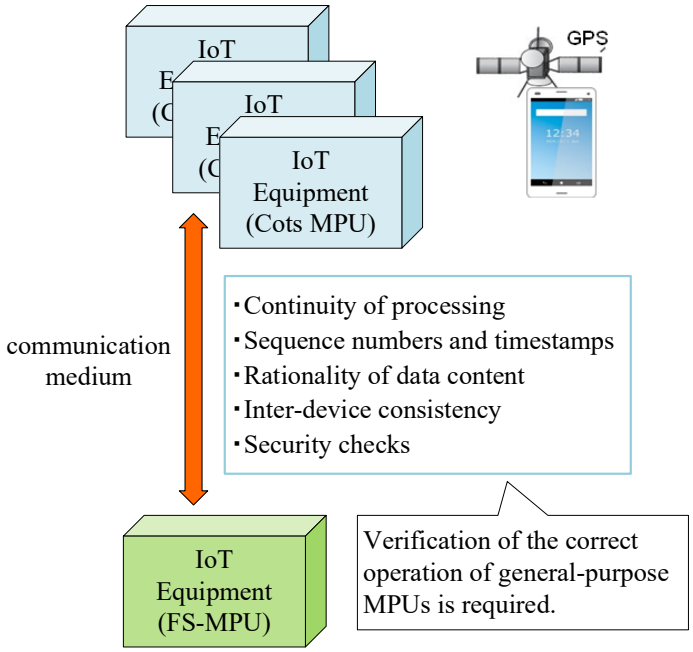


Figure 6: An example of a safety assurance method for general-purpose products

In addition to the fail-safe approach discussed in this paper, a topic of growing interest in system design is "intrinsic control". Railway signalling safety systems consist of station interlocking systems and block control systems, which together ensure the safe operation of trains. Recently, wireless train control systems have emerged as the latest development in this field. In such systems, trackside train detection devices and wayside signals are not used. Instead, safety is ensured by issuing a Limit of Movement Authority (LMA) from the ground system, which governs safe train operation. In intrinsic control, conventional interlocking and block control systems are eliminated. Instead, essential information is exchanged among points, level crossings, trains, and other system components, and verified safe conditions are transmitted to trains as LMA. Figure 7 illustrates the system configuration for intrinsic control, while Figure 8 shows the control method [10]. To implement this system, it is essential to ensure interoperability through standardized specifications and to establish a multi-vendor environment that enables broad participation by manufacturers. In this context, the "Study Committee on Train Control Systems for the Introduction of Autonomous Driving on Regional Railways"

has been established under the Ministry of Land, Infrastructure, Transport and Tourism, and has been examining this topic since November 2024.

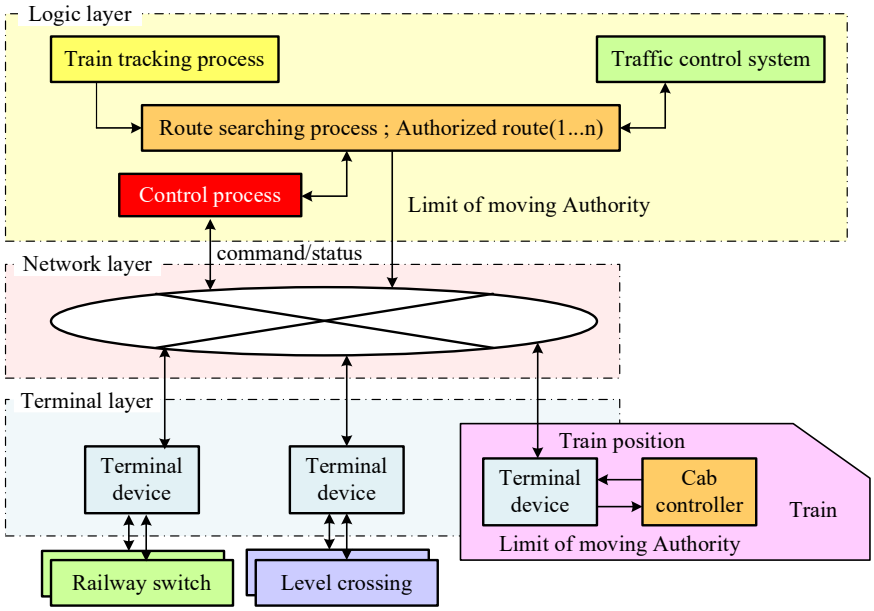
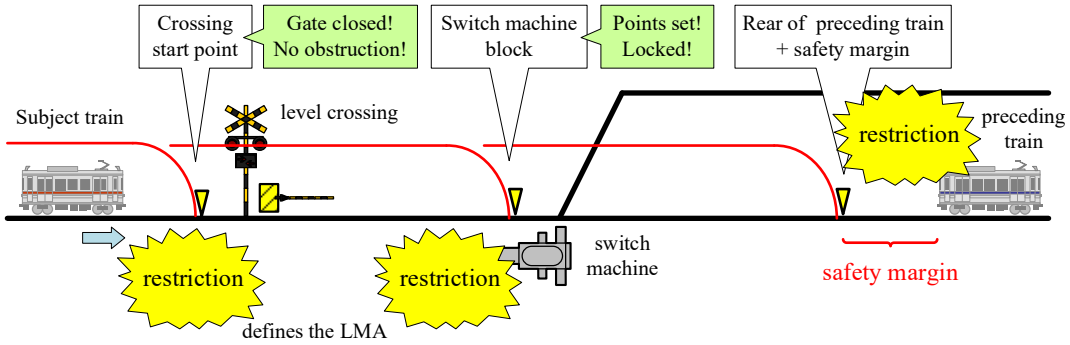


Figure 7: Concept of a new train control system based on intrinsic control



Identify the nearest LMA ahead.
A level crossing ceases to define the LMA when barriers are closed and no obstruction is present. Points cease to define the LMA when set and locked for the train, allowing the search to extend further ahead. A preceding train defines the LMA up to its rear plus a safety margin. The train may proceed up to the LMA.

Figure 8: Train control method based on movement authority（LMA）

3 Applicability of Generative AI for Signalling Control

As the focus of safety assurance shifts from fail-safe hardware design to system-level integrity verification of inter-component interfaces, the complexity and amount of information to be evaluated increases significantly. Traditional deterministic logic design and manual verification methods are difficult to adequately address this

complexity. Therefore, new methods that can evaluate the validity of complex data sequences are needed, and the application of generative AI is expected to be a promising candidate.

Generative AI is advancing rapidly, and there are high expectations that incorporating this technology into signalling control systems will improve functionality and reduce costs. However, due to issues such as hallucination in generative AI, its adoption remains cautious at present. In this chapter, we present this perspective and discuss the concept of applying generative AI to railway signalling control systems to address these challenges.

3.1 Challenges in Applying AI

The Institute of Electrical Engineers of Japan (IEEJ) formed an expert committee on the application of advanced safety technologies and risk management methods to railway signalling to discuss the introduction of AI technology to railway signalling systems. The report was compiled in December 2023. In light of these discussions, the challenges of applying AI technologies, including Machine Learning (ML) components, to safety-critical control systems such as railway signalling control systems can be summarized as follows.

(1) Inability to fully define input–output relationships in ML components
ML components operate based on learned data; however, their behavior cannot be fully specified when inputs fall outside the training dataset. As a result, the safety of decisions made by ML components in safety-critical control systems cannot be guaranteed.

(2) Limited explainability of ML decision-making
ML components generate outputs from inputs, but the underlying reasoning is often difficult to interpret. If an unexpected output occurs during operation, operators cannot reliably determine whether to continue system operation. Furthermore, when failures occur due to incorrect decisions, it becomes difficult to identify corrective actions if the reasoning behind the decision remains unclear.

3.2 Proposed Approach

Because of these difficulties, there are significant concerns regarding the use of generative AI in safety-critical control systems such as railway signalling. To address these issues, we propose an alternative approach in which generative AI is not used for control decisions. Instead, its capability for data exploration is utilized to evaluate the consistency of input–output sequences at system interfaces. As illustrated in Figure 6, safety in intrinsic control is achieved not by centralized control logic, but by consistency verification at the interfaces between components. In this framework, generative AI evaluates whether the sequence of input–output pairs at these interfaces is reasonable, based on the processing logic and rich control data accumulated in existing systems. By restricting the role of generative AI to consistency evaluation rather than direct control, this approach leverages its strengths while avoiding its

inherent limitations. A prototype system is currently under development, and the results will be reported in a future study.

4 Conclusions

In this paper, we have shown how the concept of fail-safe design has evolved and have outlined a new proposal for the application of Generative AI to signalling control systems.

First, we reviewed the development of fail-safe design technologies embedded in signalling systems. In particular, we discussed fail-safe relay circuits with asymmetric error characteristics, the development of fail-safe logic systems, and their relationship with fault-tolerant technologies.

Furthermore, focusing on the fact that many accidents are caused by inconsistencies at system interfaces, we argued for a shift toward simplified system architectures with fewer components. In this context, we introduced the concept of intrinsic control, in which systems are realized through the exchange of only essential information. Under intrinsic control, conventional interlocking and block control systems are eliminated.

Finally, regarding the use of Generative AI, we identified key concerns and proposed an approach to address them. Instead of delegating control decisions to Generative AI, our method utilizes its capability to evaluate data consistency while maintaining safety. A prototype system is currently under development, and the results will be reported in a future publication.

Acknowledgements

The authors would like to express their sincere gratitude to Mr. Takahiko Ogino (formerly of the Railway Technical Research Institute) and all those involved for their valuable guidance, advice, and support throughout this study.

References

[1] Watanabe and Takahashi, "Generalization of Fail-Safe Logic Systems", IEICE National conference, 72, (1965-11) (in Japanese).
[2] Okumura and Watanabe, "A Study on Interlocking Devices for Railway Signals Using a Fail-Safe Electronic Computer", Railway Research Report, No. 803, (1972) (in Japanese).
[3] Toma, Oyama, and Sakai, "A Construct Method of Fail-Safe Sequential Circuits Considering Symmetry Errors", IEICE (D), 55-D, 3, pp.202-209, (1972-3) (in Japanese).
[4] Tsuchiya, "Three-value C-type fail-safe logic circuits", Kejigakuron, 6,1, pp.81-88, (1967-2) (in Japanese).
[5] W.C.Carter, et al.: “Cost effectiveness of self-checking computer design”, Dig. Pap., FTCS-7, pp117-123 June 1977.

[6] Nanya and Kawamura, "The Concepts of Error Safety and Error Propagation in Self-Checking Systems", IEICE Theory, J68-D, No.12, pp2007-2014, (1985) (in Japanese)
[7] Ohno and Tsunoyama, "A Method for Fail-Safe Configuration by Microcontroller", IEEJ Research Report, PAT-83, 16, pp.49-58 (1983) (in Japanese).
[8] B.J.Sterner, ”Computerized Interlocking System ” , Railway Engineers International, Vol.3, No.6(1978).
[9] IEC61508-3: Functional Safety for Electrical, Electronic, Programmable Electronic (2010).
[10] Yoshihisa Saito, Akira Asano, Hideo Nakamura, Hiroshi Mochizuki, Hijiri Takahashi: "Reconstruction of Train Control System by Hierarchy", IEEJ Expert Committee on the Application of Advanced Safety Technologies and Risk Management Methods to Railway Signals, ed.: "Investigation on the Application of Advanced Safety Technologies and Risk Management Methods to Railway Signals," IEEJ Technical Report (2023.12) (in Japanese) .