Can Artificial Intelligence (AI) in Robot Systems Manage the Functional Safety Parts?

Innovation in robotics and sensing technologies have led to increasingly wide use of production systems whereby humans work in collaboration with robots at industrial manufacturing sites. Recently, studies on innovations in computing resources have shown improved productivity thanks to more advanced control and intelligent movements. In one of the multiple studies, carried out so far, efforts have been made to develop robots that integrate control functions and AI to determine the tasks and to carry out autonomous controls and tasks.

When robots are used at industrial manufacturing sites, to ensure the workers safety, it is necessary to perform a risk assessment. Generally, such assessments are conducted from the perspective of Functional Safety. Since robots that have control functions and embedded AI, we intend to prove the facts found during our evaluations.

As a result, it was found out that the facts encountered due to integrating AI were not easy to solve. Therefore, approaches for improving the advantages of production systems that use collaborative robots and AI should be restricted in very limited areas.

Main Problems and Issues

Generally, the elements employed by robots to determine autonomous control actions are classified into five main topics:

- Environment recognition
- Action planning (task planning and path planning)
- Trajectory generation
- Motion control
- Measurement

Regarding industrial robots, in order to use AI in an integrated manner with these elements, Functional Safety must be evaluated in detail. The topics evaluating Functional Safety were examined from two perspectives:

- The difficulty of judging the states of AI systems
- The differences in performance evaluation metrics between Functional Safety and AI systems

This document is based on the results of those evaluation.

WHITEPAPER

1. The difficulty of judging the states of AI systems

It is difficult to distinguish whether AI systems are in a normal state or in a failure state. This can be shown by considering a case in which the state of a redundant AI system is judged by comparing its Input and Output states.

As shown in the figure below (Figure 1) , supposing that if both Inputs and Outputs are the same between the redundant AI systems, it is judged to be normal state, while if the Inputs are the same but the Outputs are not the same, it will be judged to be faulty.

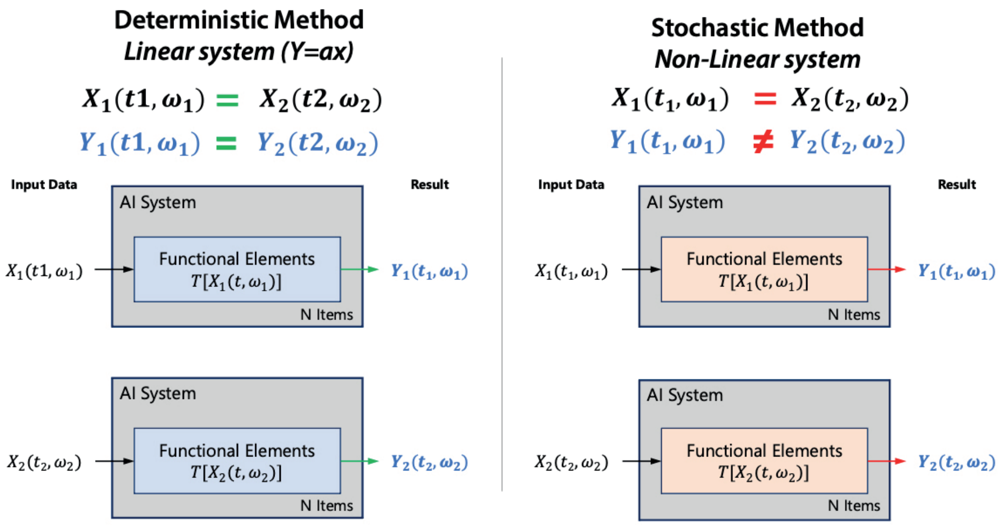


Figure 1 Comparison of Deterministic and Stochastic (Random) Methods

A deterministic linear method has always the same results irrespective of the time changes, which means that independence of the time or of the number of repetitions, the results will remain the same. Thus, in a Stochastic non-linear method, with the same inputs and same operation Functional Elements, each time the results could be equal or different.

According to the architectures in Figure 2 of **ensemble learning** and considering how the result will be affected if the bagging and boosting structure tree fails: If this is applied to the k-out-of-n architecture for Functional Safety, bagging makes a redundant parallel tree architecture, while boosting makes a serial single-channel tree architecture.

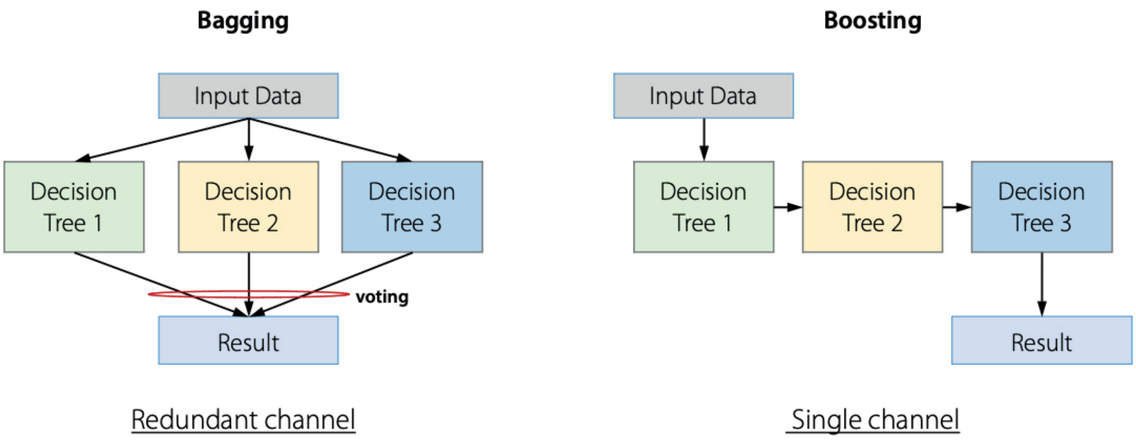


Figure 2 Architecture of Ensemble Learning from the perspective of Functional Safety

Bagging seems to have an advantage in front of Boosting from the viewpoint of redundancy, but it cannot guarantee that the final results are the same as the previous fault.

Therefore, if the system is normal cannot be judged, and it cannot be said to have high faulty resistance.

For neuronal networks (as exemplified by deep learning), which have a complex architecture of nodes or layers, diagnosing their judgment results is not a common practice; thus, if the system is normal, it will be judged based on the final Output result. See Figure 3 below.

However, since it is difficult to find faults in some hidden layers and elsewhere, it cannot be judged if the system is in normal state.

These examination results provide an explanation why AI systems cannot be treated as deterministic systems for the purposes of Functional Safety.

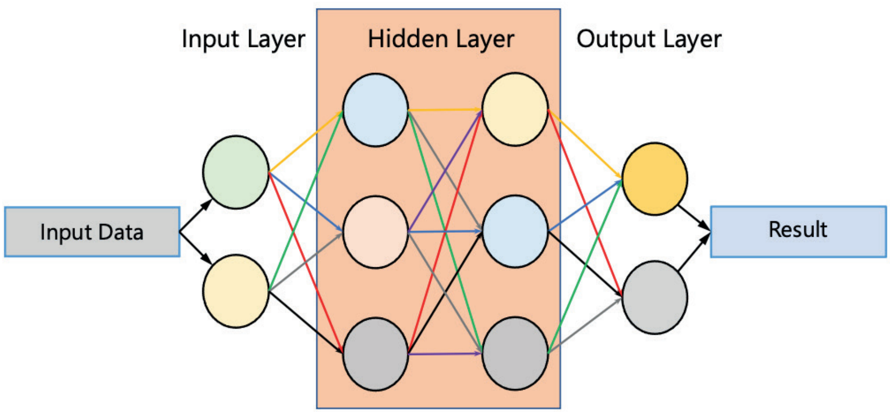
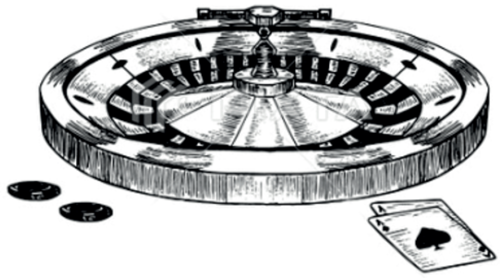


Figure 3 Architecture of neuronal network from the perspective of Functional Safety

To sum up, since **stochastic decisions** are made when AI systems fail, a robot system with AI incorporated into its autonomous control cannot be evaluated as a deterministic system.

Therefore, such a system cannot be evaluated as part of the Safety-Related Parts of the Control System of Functional Safety, which must be evaluated as a deterministic system.



2. Differences in performance evaluation metrics between Functional Safety and AI systems

Functional Safety is evaluated by the assumption that the failure probability during on-site operation (during the chance failure period) is constant.

On the contrary, AI systems can produce normal results for data during operation that could be like the learning data used during learning and that is assumed to be steady, but it is not guaranteed that data like the data learned during on-site operation will be an Input.

Therefore, the error probability of AI systems changes from time to time, and it cannot be evaluated by the assumption that the probability of failures during on-site operation are constant as it is done for Functional Safety.

Performance evaluation metrics for AI include accuracy, precision, recall, and specificity. AI depends on learning data that does not take time series variations into account; thus, if these performance metrics are evaluated using a confusion matrix (Figure 4) for test data to solve a binary classification problem, evidence of effectiveness cannot be presented from the perspective of Functional Safety.

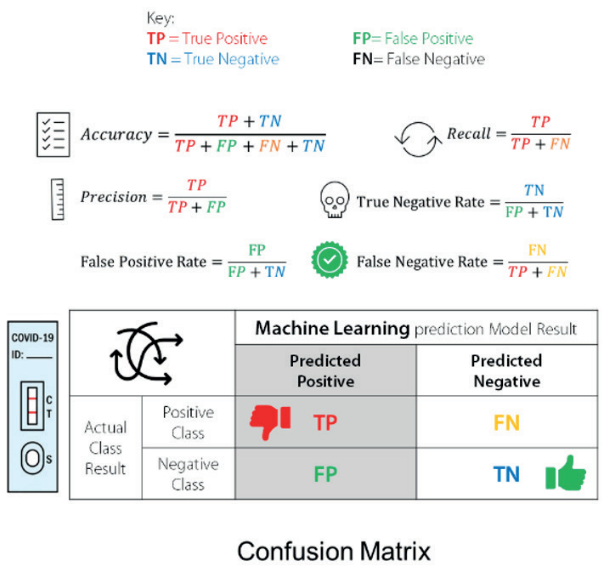


Figure 4 Confusion Matrix

Actually, if a machine learning model that has produced good results from test data is evaluated on the assumption with time-varying test data, then it is not possible to guarantee the result to obtain the same accuracy. This means that a model that has produced good results from test data is only a single sample at one point in time.

The performance of AI systems with no time-varying parameters cannot be evaluated from the perspective of Functional Safety. Consequently, their safety performance cannot be evaluated without non-functional-safety metrics.

Any of these issues evaluating the safety of AI systems are not easy to solve.

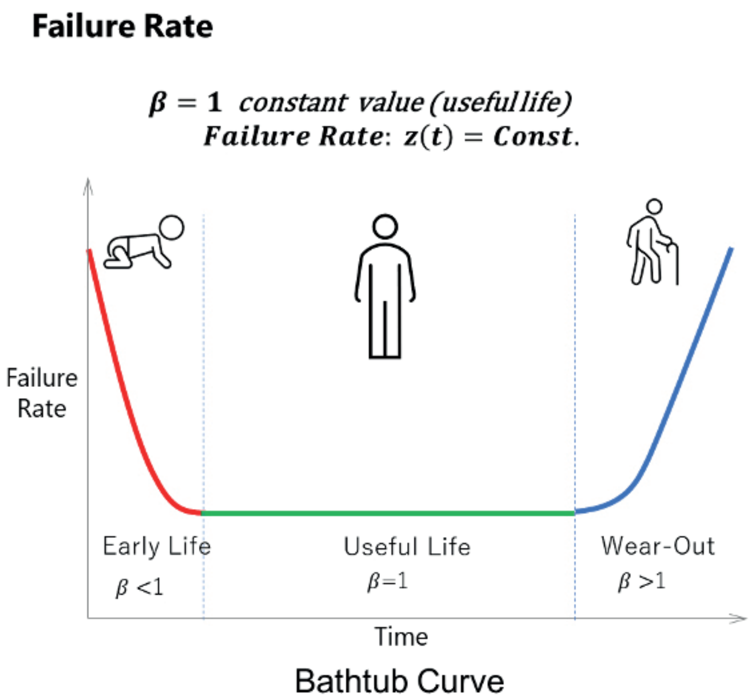


Figure 5 Functional Safety failure rate

In conclusion, today it is useless from the perspective of Functional Safety to develop industrial robots with AI integrated into their control systems. Thus, AI can only be used in areas where its characteristics can improve the production and productivity of non-safety systems. At the same time, we are waiting for the coming generations of computing system improvements as well as new additional features.

Authors:

About Atsushi Oshiro

Mr. Oshiro is an experienced Engineer, specialized in AI, Probabilistic Robotics, Robotics Control Engineering and Functional Safety.
He is currently studying for a PhD at Japan Advanced Institute of Science and Technology in Nomi, Japan. One of his key focus areas in engineering is Artificial Intelligence. Currently, Mr. Oshiro is a Manager at the Development Center in Kyoto at OMRON Corporation, Japan.

About Josep Plassa

Mr. Plassa is a seasoned Safety Expert with more than 25 years of experience in Safety Automation. He is also a Safety Independent Inspector accredited by Deutscher Akkreditierungs Rat (Germany). Nowadays he is a permanent member of Standardization Committees ISO and IEC, such as Industrial Robotics Safety, Machinery and Safe Motion, representing Spain-UNE. Mr. Plassa is currently a Safety Product Marketing Manager at OMRON Europe, based in Barcelona