Poly-agent systems approach to organizational accidents: Theory and its applications

Masuda Hiroyuki and Kijima Kyoichi

Tokyo Institute of Technology, VALDES (Dept. of Value and Decision Science),

Graduate School of Decision Science and Technology

h-masuda@valdes.titech.ac.jp , kijima@valdes.titech.ac.jp

Abstract

The purpose of this research is to analyze and obtain insights for preventing fatal accidents in organization-based projects by poly-agent systems approach, a kind of systems thinking. To prevent such accidents, it is certainly indispensable to recognize and understand potential danger at the stages of "concept making" and "predetermined definition" of the project management. Based on the poly-agent systems approach, this paper points out a crucial factor of organizational accident inconsistencies among the subjective understandings of the agents (or decision makers) involved in the situation. To achieve the safety culture, we examine the learning process in the organization, since we believe learning is the core process of safety management. In this paper, we propose the new system reliability and safety paradigm called "Reliability Analysis & Synthesis based on the agent system theory (Agent-Based Reliability Theory)". We also agree the validity of our scheme by examining some actual cases.

Key Words & Phrases : Organizational Accidents, Safety, Poly-Agent Systems Approach, System Design, Accident Prevention, Human Error, Risk Management, Symmetric Information, Asymmetric Information

1. Introduction

The purpose of this research is to analyze fatal accidents in organization-based projects to obtain insights for preventing them by poly-agent systems approach [7], a kind of systems thinking. J.Reason [1,2,3], who firstly proposes the concept, defines organizational accidents as comparatively rare but often as catastrophic events occurring within complex modern technologies. They may legitimately be seen as products of technological innovations that have radically altered the relationship between systems and their human factors. For example, the JCO accident, which occurred at a nuclear fuel conversion facility, Tokai-mura, Japan on 30th Sept. 1999, was a typical and serious organizational accident.

It is absolutely true that interdisciplinary risk management approaches are necessary to tackle safety problems concerning organizational accidents. However, each discipline takes its own viewpoint: For example, the field on project management have discussed it indispensable to recognize and understand potential danger at the stages of concept making and predetermined definition of the project. They also argue that diagnosis for accidents prevention is necessary in system design phase. In engineering approach, on the other hand, they have excluded the failures caused by intentional actions such as sending incorrect reports. The human reliability analysis assumes human errors are never blamed since rationality of human being is bounded [4,15]. It claims that we should not impose sever penalty on those who have made errors in order to encourage, rather than blame, to tell why such accidents happened honestly for preventing similar accidents to happen again in the future.

We believe that we definitely need not only engineering or "hard" approaches but also human/social scientific viewpoints to prevent human / organization errors, because fatal human errors are often caused by complex mixture of many factors. The implementation theory in economics, for instance, which discusses design of incentive mechanisms that motivate the decision makers to tell the truth, may refer to this problem. It is because built-in of some kinds of safety mechanisms in a society certainly helps to discourage people to make such errors intentionally with malice that cause serious accidents.

In this paper by adopting poly-agent systems approach, a contribution from the systems thinking discipline, we try to examine why such fatal accidents happen and how we can prevent them. The poly-agent systems approach is proposed to analyze and investigate behavior of decision makers (or agents) and their interaction of societal systems. It explicitly assumes that the agents make decisions autonomously based on their subjective understanding of the world. Furthermore, it assumes that tight interaction among the agents through networks, the information flow, either symmetric or asymmetric, give influence on and their decision making process.

Based on it this paper points out a crucial factor of organizational accident consistencies among the subjective understandings of the world (or the environment) of the agents (or decision makers) involved in the situation. That is, since each agent may see even the same situation differently, their misunderstandings or misperceptions most likely lead to fatal accidents even though each does his/her best. We also agree the validity of our scheme by examining some actual cases.

2 Learning: A Crucial Activity for preventing organizational accidents.

Though Reason published the concept of organizational accidents [1], but he did not give clear taxonomies of the accidents. Fig. 1 shows taxonomy of organizational accidents derived from the discussions above. It classifies organizational accidents into two categories; *i.e.,* those that occurs accidentally or unintentionally and those that occur intentionally.

We claim that to deal with organizational accidents, we need to take into account human/social science-based approaches called poly-agent systems approach, because reliability engineering , that is traditional engineering approach for the accidents, is not sufficient to prevent organizational accident , so that more "soft" concepts such as trust and incentive design should be necessary.

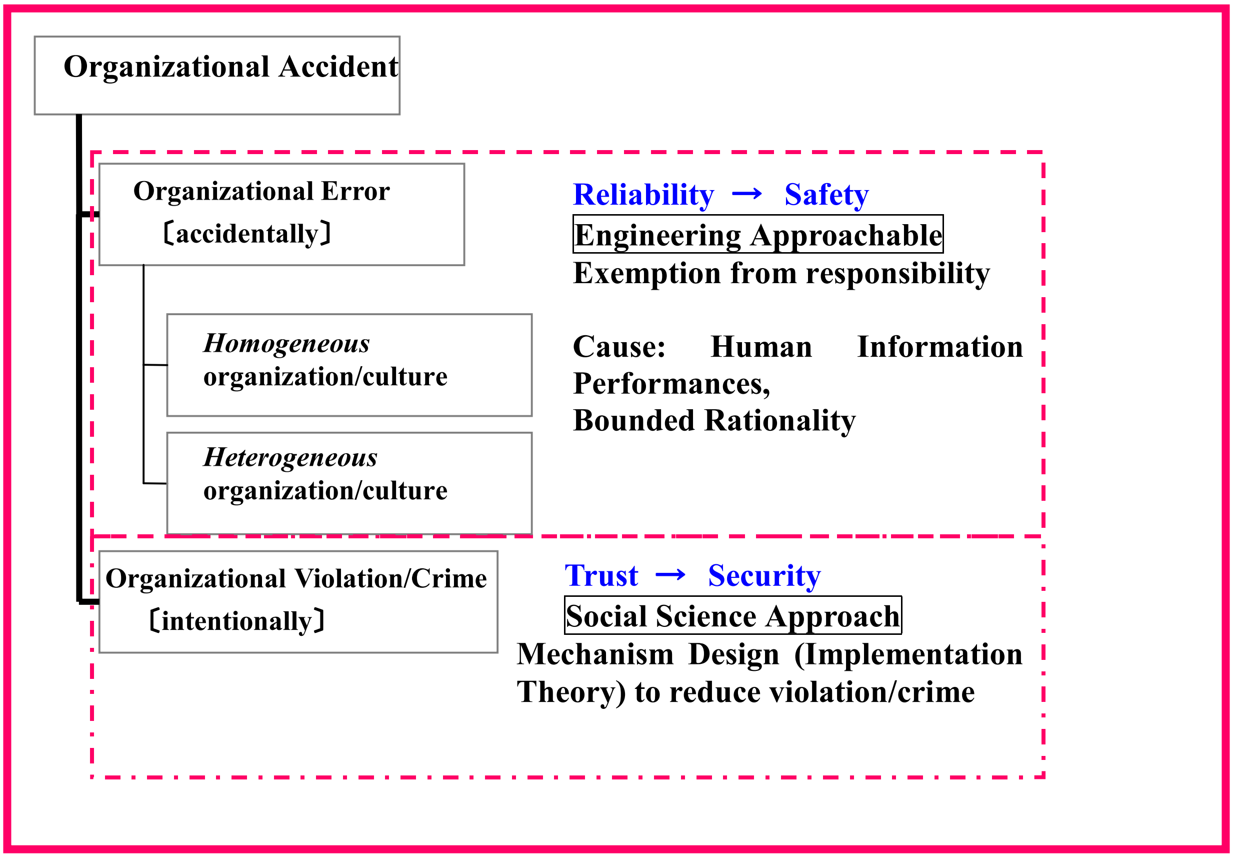

Fig.1 Structure of Organizational Accident

By analyzing well-known organizational accidents so far, it was pointed out that they were fundamentally caused by failures of communications among the agents. For example, for medical accidents: it is often claimed that "small accidents happened all the time, but they were not reported , but large accidents stem from collection of small ones. So it is important to learn small accidents especially among doctors/nurses." [19,22], the comment from Atomic Society of Japan on the JCO accident says "The workers' unsafe actions that directly caused the accident are different from the basic human error types, that is, slip, lapse, and mistake, because their actions intentionally deviated from the normative procedure. At the same time, they are clearly different from sabotage, because the workers instead of expecting bad outcomes, they were motivated to high productivity, though they might have been aware of illegitimacy of the procedure."[20]

The communication gaps generate and increase inconsistency or discrepancy among the understandings of the world of the agents. In the case of medical accident, doctor/nurse to patient or doctor to nurse communication is the most important to prevent the medical accidents, that is, risk management in the hospital [5,6,19].

Learning from failure is perhaps the most effective way to prevent organizational accidents. Indeed, Fortune [9] emphasizes not only analysis but also synthesis in her particular technique supporting the agents to learn from failure (Refer to Fig. 2). She especially said, "In the case of the Systems Failure Method, the goal is a systemic interpretation of a failure and its context which could in turn lead to some action." and she also say that " the Systems Failure Method has two key features: conceptualization and modeling of the failure situation as a system(s); and comparison of that system(s), first with a model of a robust system that is capable of purposeful activity without failure, and subsequently with other models based on typical failures. Its full diagrammatic representation is shown in Fig. 2".

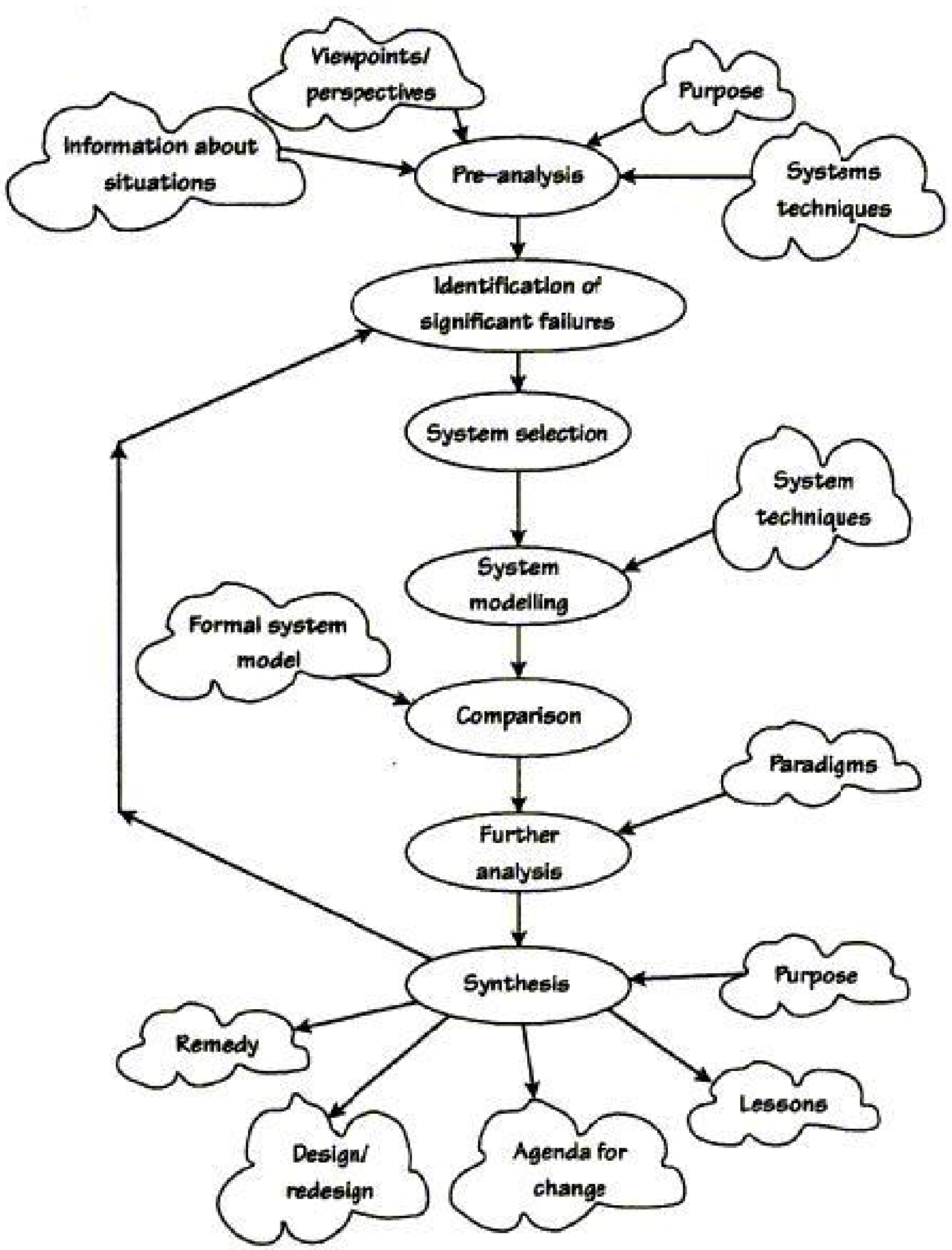

Fig.2 The Systems Failures Method (SOURCE: LEARNING from FAILURE)

Her scheme for learning is quite useful to understand accident analysis concept. However, the perspective of organizational accidents is not so clear. The organizational learning certainly requires mutual understanding of the agents and their proper communication, because learning is essentially reciprocal activity. Furthermore, as we argued above, organizational accidents are considered, especially if they are unintentional, as negative emergent events due to malfunction of the interactions in the system [1].

These arguments naturally lead us to focus on the relationship or interaction among the agents when analyzing organizational accidents. The poly-agent systems approach is an approach that discusses these aspects explicitly.

3. Models of Communication in Organization

3.1 Three types of relationship

To understand the occurrences of failure or accidents in organization, we need to consider some factors, that is, agent(s) , their relationships or actions, and environment. First we model the features and relationships between agents and environment/field as following three types.

1.Nesting Structure of Agents

Agents have a distinctive feature that they may get together autonomously to form groups or organizations. In these organizations, they may well make sub groups or teams, we call the sub structures sub agents.

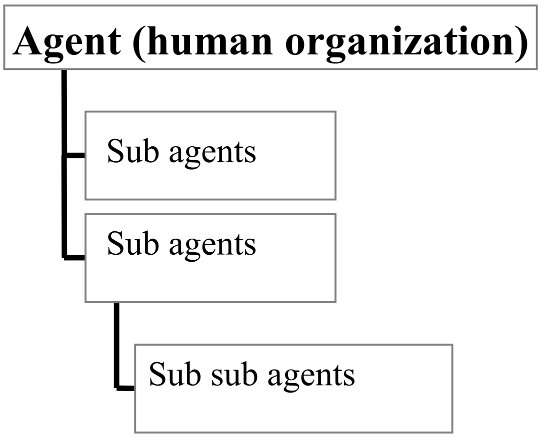

Fig.3 Nesting structure of agents

2. Agents to Environment/Field Interaction

The agents can interact with the environment/field by giving and taking information and other resources to and from the environment. The environment may include artificial environment (or articrafts), social environment, and natural environment.

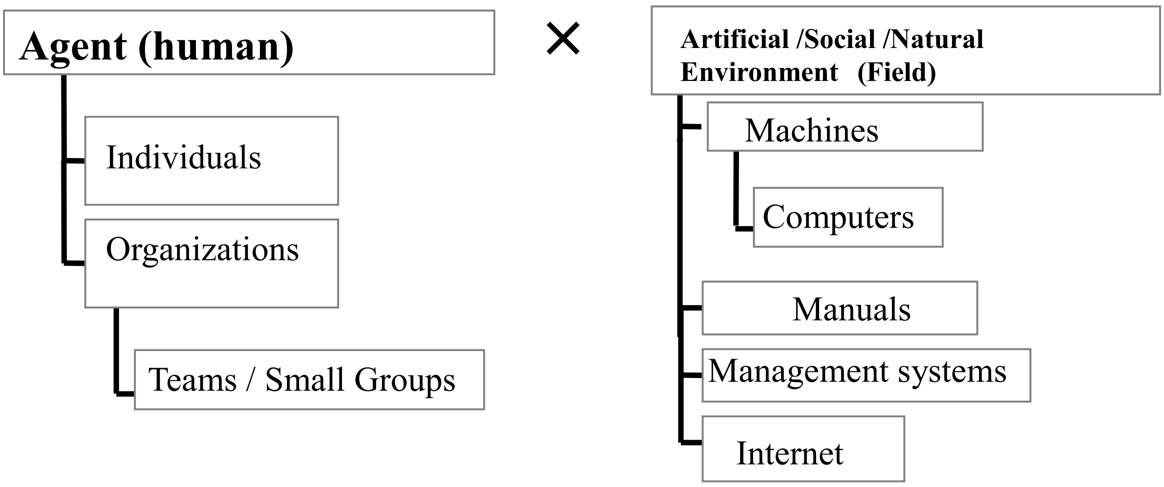

Fig.4 Agent × Environment Interaction

3. Agent to Agent Interaction

The agents interact with each other autonomously by exchanging the information by communication. Learning requires an appropriate communication that exchanges knowledge among agents.

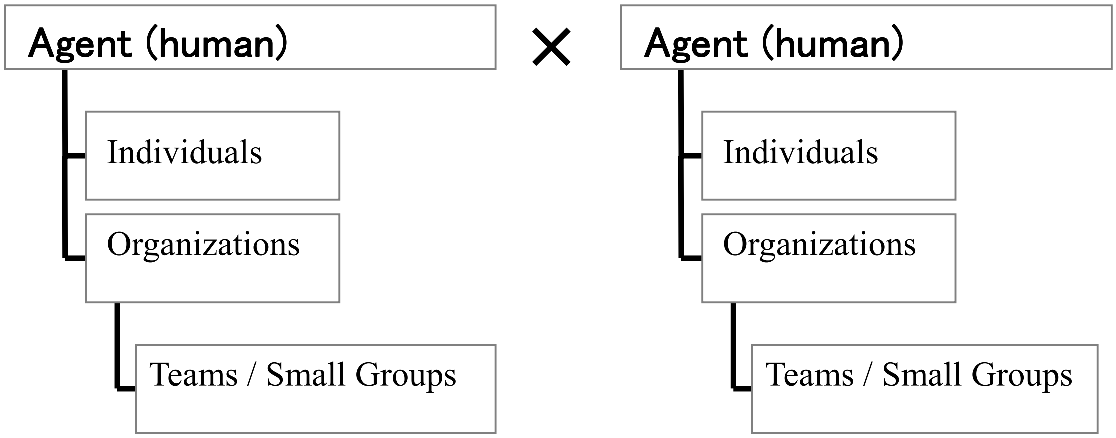

Fig. 5 Agent × Agent Interaction

3.2 Symbolic Model Representation

The structure of symmetric or asymmetric information flow plays important roles in decision-making process. Basic principle teaches that the organization should clarify the role and the responsibility of each worker, *i.e.,* clarification of authority, responsibility, and accountability. [10].

In the model of this research, we assume the agent plays a role either as a decision maker or as an actor. If the agent has authority then he/she becomes a decision maker with rights to instruct other agents to do something. Therefore the agent (decision maker) also has responsibility. The actors who don't have the authority follow the decision maker. On the organization accident, when we pay attention to the relation, that is, communication between the decision makers and actors often fails. Mainly, this originates from one-sided information, so called, from asymmetric information. It is considered as lack of accountability, or reports from actor to decision maker. Such a situation is shown in Fig.6, as "One Way Communication Model".

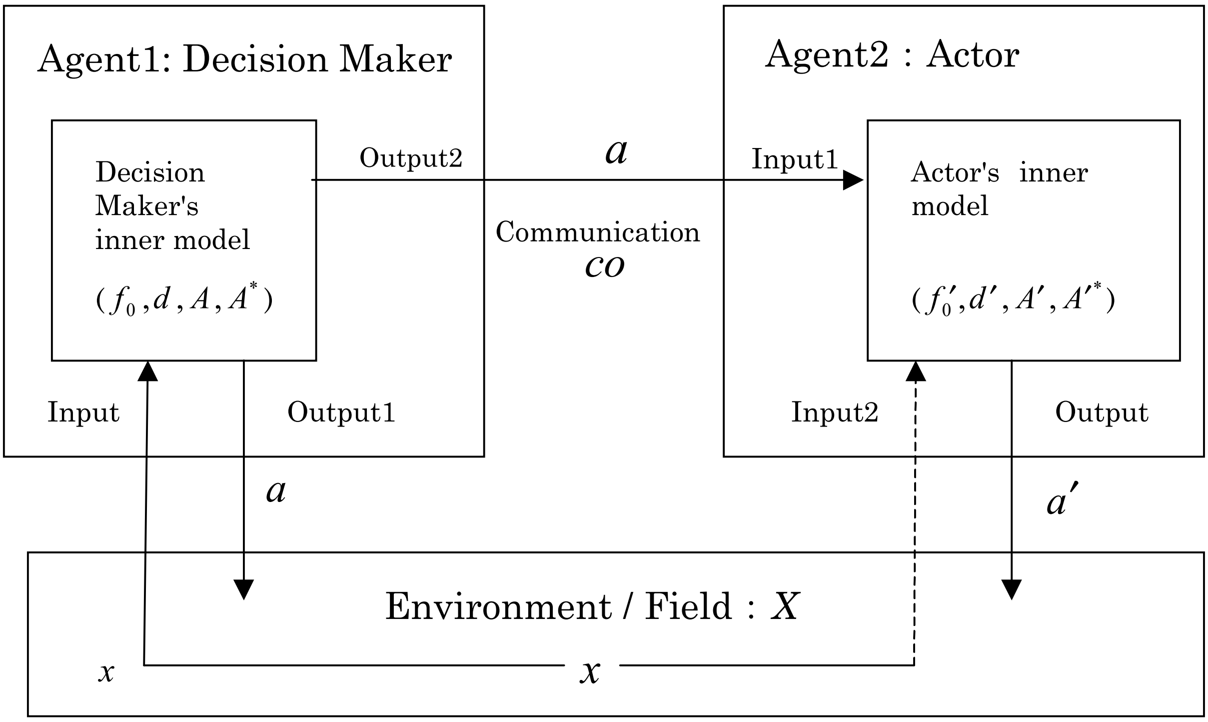

Fig.6. The framework of One Way Communication Model
(Lack of intercommunication)

X is the set of environment / field, and it is assumed that $x \in X$ occurred. Now we call f_0 and f_0' decision makers' observation function and actors' observation function, respectively. A and A' denote decision makers' action set and actors' action set, respectively; $A^* \subset A$ and $A'^* \subset A'$ represent decision makers' ERROR action set and actors' ERROR action set. While, d and d' denote decision makers' decision functions and actors' decision function, respectively. Then, we call quadruples (f_0, d, A, A^*) and (f_0', d', A', A'^*) decision makers' inner model and actors' inner model (See Fig.6). Decision makers observe $x \in X$ and perceive it as $f_0(x)$. Using decision function d, he / she chooses an action $a = d(f_0(x)) \in A$. On the other hand, the actors select an action $a' = d'(CO(a), f_0'(x)) \in A'$, based on information $CO(a)$ given by the decision-maker and $f_0'(x)$ observed by the actor, where $CO : A \to A'$ is an information transfer function. If the structure of d and d' are different, that situation involves asymmetry information.

We identify possible errors that may happen in the one way communication as follows:

1. Decision-maker's action $a^* \in A^*$ is an error action, so that he/she did an error action through Output 1.
2. The decision maker's action $a \notin A^*$ is a correct one (NOT an error action), but through communication CO, the information is not transmitted from the decision maker to the actor correctly; or the information is not received by the actor.
3. Though $a \notin A^*$ is correctly perceived by the actor, a wrong decision making $a'^* \in A^*$ is done by the actor because he/she has inner model different from the decision maker's understanding.

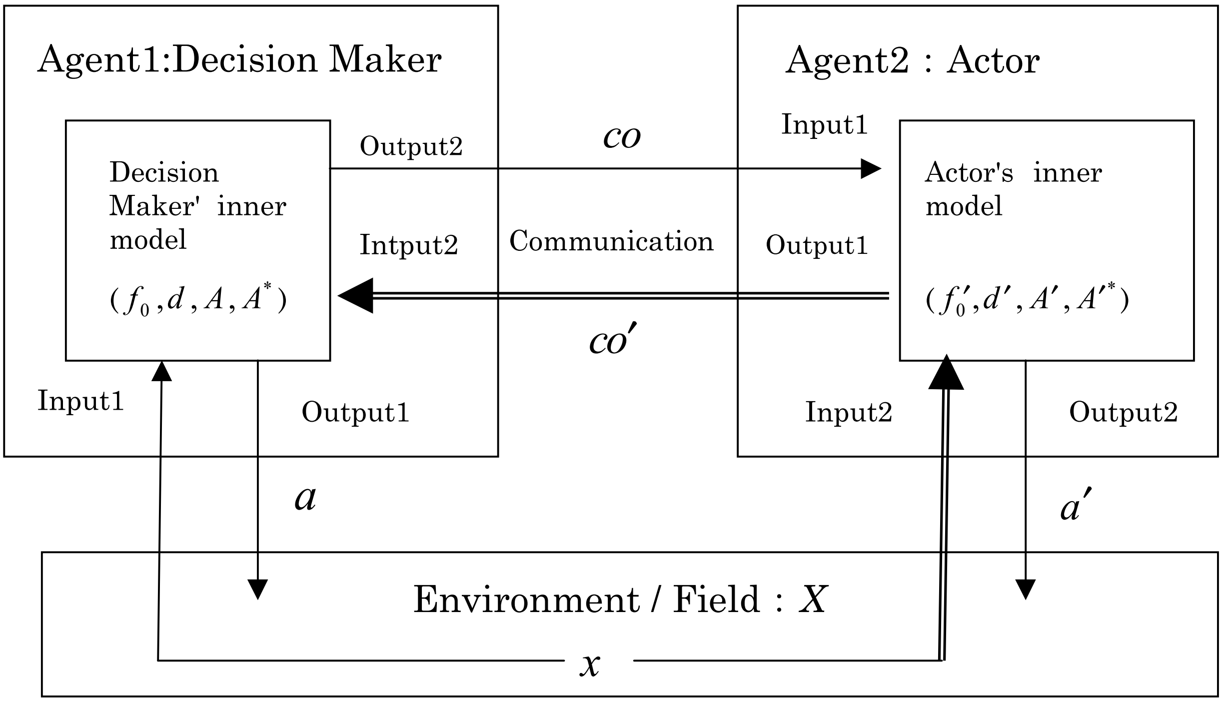

Fig.7. The framework of Intercommunication Model

Based on the weakness of the one way communication model, we insist that the framework of model described by Fig.7 as a preferable model in this paper. Compared with the one way communication model, the advantages of the intercommunication model in Fig.7 are:

1. actors observe themselves more autonomously (voluntarily) and the anticipation to the errors of agents is improved and,
2. close communication happens between the agents.

In Fig.7, if $a = a'$ holds, it means that there is mutual check or mutual understanding between decision maker and actor, so it is suggested to execute it as a decision.

$$\begin{cases} a = d\,(co'(a'), f_0(x)) \\ a' = d'\,(co(a), f_0'(x)) \\ a = a' \end{cases}$$

Accordingly, learning in this paper is to adjust how to clarify agent's inner models, the decision making functions, and the information transfer functions through intercommunications, by which there is "common understanding / knowledge". That is, the focus is applied to sharing information by the learning process of the mutual recognition and communications between agents.

4. Conclusions : to achieve the safety culture

Safety control and risk management become more important in the complex and high technology society. Though organizational factors and safety culture have been argued over the world after the Space-shuttle accident and the Chernobyl accident, little attention has been paid to these factors in Japan. To achieve the safety culture, we need to examine the learning process in the organization, since we believe that learning is the core process of safety management. In this paper, we proposed the new system reliability and safety paradigm called "Reliability Analysis & Synthesis based on the agent system theory (Agent-Based Reliability Theory)". Demonstration of this concept model by computer simulation is under the way.

References

[1] Reason, James, Managing the Risks of Organizational Accidents. Dartmouth.1997

[2] Reason, James, Human Error, Cambridge University Press, 1990

[3] Kunihide Sasou, James Reason: "errors: definition and taxonomy", Reliability Engineering and System Safety, 1999

[4] Jose M Nieves and Andrew P Sage, "Human and Organizational Error as a Basis for Process Reengineering: With Applications to Systems Integration Planning and Marketing", IEEE TRANS. SYST. MAN. CYBERN. vol.28, NO6, NOVEMBER, 1998

[5] Charles Vincent, Maeve Ennis, Robert J.Audley (Editor), Medical Accidents, Oxford U.P.,1993

[6] Bogner, Marilyn Sue., Human Error in Medicine.LEA,1994

[7] Kenji Tanaka, "Poly-Agent System Design Approach To Autonomous Anticipative Management", Proceedings of the Fourteenth European Meeting on Cybernetics and Systems Research, pp.14-17, 1998

[8] Checkland, P.: Systems Thinking, Systems Practice. Wiley.1981

[9] Fortune, Joyce, Peters, G., "LEARNING from FAILURE; The systems approach.", Wiley, 1995

[10] Silverman, Melvin; Project Management, Wiley & Sons, Inc, 1976

[11] Bradshaw, J.M.; Software Agents., AAAI Press / The MIT press, pp.3-46, 1997

[12]Rasmussen, J., "Skills, rules, knowledge: signals, signs and symbols and other distinctions in human performance models", IEEE TRANS. SYST., MAN., CYBERN., SMC-13, 257-267, 1983

[13] S Beer, DIAGNOSING THE SYSTEM for organizations, John & Wiley , 1984

[14]Niklas Luhmann,VERTRAUEN,FERDINAND ENKE VERLAG,1973

[15]Simon, Herbert A.; Sciences of the Artificial the 3rd ed.,MIT Press,1996

[16]Minsky, M.:The Society of Mind.,Simon & Schuster, 1988

[17]Russell, Stuart J.,Norvig, Peter: Artificial Intelligence: A Modern Approach. (Prentice Hall Series in Artificial Intelligence) , Prentice-Hall Intl Edns., 1995

[18]Flood, R.L., "Rethinking the Fifth Discipline: Learning Within the Unknowable." Routledge, 1999

[19]The Japan Society for Quality Control: Symposium on Prevention-Its Technology and Management- abstract (in Japanese)

[20]Statement on JCO Criticality Accident:Human-Machine Systems Division, Atomic Energy Society of Japan :JCO Accident Special Work Group, http://mirage.qse.tohoku.ac.jp/aesj-mms/joc-e.htm

[21]American Heritage Dictionary, the third edition, Houghton Mifflin, 1994

[22] http://www.asahi-net.or.jp/~PF4M-ATM/IRYOU/IRYOUJIKO/990320risk.html (in Japanese)