Trans. of the Society of Instrument and Control Engineers
Vol.E–2, No.1, 150/159 (2002)

Self-referential Structure in Collective Agent System

— II Carrier Sequence Control of AGV Transportation System based on Diversity–regulation of Strategy —

Toshimitsu HIGASHI*, Kosuke SEKIYAMA** and Toshio FUKUDA***

This paper proposes a system, which realizing a collective autonomous behavior such as an autonomous conveyance order formation in the AGV(Auto Guided Vehicle) transportation system. We attempt to deal with a large scale distributed autonomous system in dynamic environment feasibly. We have worked to realize the dynamically reconfigurable formation in the dynamic environment, and we showed that the dynamically reconfigurable formation emerges as the autonomous conveyance order formation of AGV transportation system in the dynamic environment. But It could not be mentioned when the restoration of the unloading success probability was caused by the self-organization, since the index of the self-organization is not clarified. In this paper, We define a strategy diversity as an index to the self-organization, and it is shown that it carries out the restoration of the unloading success probability with the diversity-regulation process of the strategy. Moreover, it is shown that adaptable parameters to the dynamic environment can be found by a distribution of "Self-organizing velocity", which is difference of strategic diversity.

Key Words: Distributed Autonomous Robotic System, Self-organizing Manufacturing System, AGV Transportation System, Collective Autonomy

1. Introduction

Emergence of collective autonomous behavior in DARS (Distributed Autonomous Robotics Systems) is expected to have wide applicability in the field of engineering as well due to its diversity, flexibility and adaptability. Recently, production systems including material handling systems are tending to move to large scale systems with high level of controls, necessitating a response to complicated and multiple specifications. Further the need to absorb change in specifications into the system and design a fail proof system that would not stop even in the event of a fault has led to the design and installation of large-scale systems designed for continuous operation.

In order to design and set up a system that is independent of the dynamic environment like change in specifications, or breakdown of equipment, there is much anticipation regarding the emergence system design theory and Distributed autonomous design methodology for realizing collective autonomous behavior based on collective intelligence. The authors propose that the Distributed autonomous design methodology is a superior methodology due to its fault tolerance and flexibility in responding to change in specification and although it has not been possible to explicitly clarify the behavior of each individual agent in the system, collective autonomous behavior displayed by the AGV in the AGV transport system that enables the system objectives to be reached have been verified through simulation [1) 2)]. As shown in Fig.1, by using Distributed Autonomous Designing Methodology proposed in the above research, even in uncertain environment where required conveyance order is not clarified, it is possible to expect the emergence of collective behavior as the agents will create the autonomous formation of the conveyance order.

Recently, various research works have published the role of Distributed Autonomous Designing Methodology for realizing the emergence of collective intelligence in uncertain and dynamic environment, namely the emer-

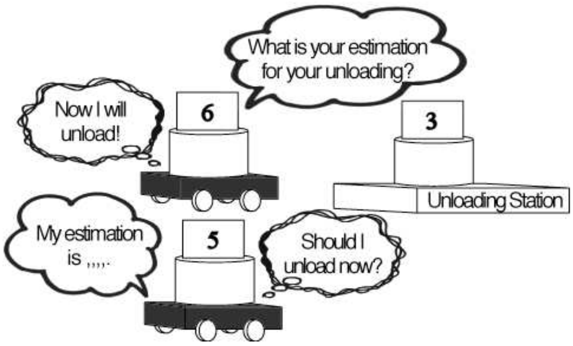

Fig. 1 Concept of collective behavior emerged as self-organizing carrier sequence

* Murata Machinery, LTD., 2, Nakajima, Hashizume, Inuyama, Aichi
** Fukui University, 3–9–1, Bunkyo, Fukui
*** Nagoya University, Furo–cho, Chikusa–ku, Nagoya, Aichi

gence of cooperative behavior through stochastic gradient methodology in the field of reinforced learning [5], Interactive schema ecosystem model for internal observation and internal description formats [11], cooperative behavior by non-homogenous robot by combination of reflecting model and planned model operations. [6], learning of multi agent behavior through fuzzy inference interaction [4] The autonomous formation of the conveyance order forming the focus of this research resembles the scheduling problem in which research covers the operation scheduling of multiple AGV [10], and emergence of cooperative behavior based on the knowledge obtained from the simulation of group behavior [8].

As shown in fig. 1, we have proposed a Distributed Autonomous Systems to explain the emergence of collective autonomous behavior occurring in the group intelligence system even in an environment where required conveyance order is not defined clearly. In order to facilitate the collective autonomous behavior, complicated strategy formation where conveyance order is not defined is realized through the proposed distributed autonomous system based on the learning achieved through autonomous behavior and mutual operation based on relative evaluation between the agents [3]. According to these results, system objectives could be achieved through AGV strategy formulation, however, in order to verify if the strategy formulation was the result of Self organizing through collective autonomous behavior could not be done as the index of organizing was not defined, so it was not possible to provide a reference of the correlation between the collective autonomous behavior of each agent and the Self organizing of the system as a whole.

We consider that Self organizing is achieved by creating a physical time span pattern that coordinates with the environment required for achieving the objective, reducing the degree of freedom of the system and through the structure formation achieved by this process, it is possible to achieve the objective of the system as a whole. We also consider that any activity or environmental change in the system will cause the structure to dynamically collapse and reform itself, and the process of (1) self reference coupling construction, (2) relative evaluation by internal observation and (3) by expansion/contraction of strategy diversification are essential for realizing this collective intelligence system. The self reference coupling structure represents a structural formation of autonomous group formation which enables self organization of relative system status based on the relationship with the environment and regulated structure existing at the location. From the fact that internal observation of the intelligence functions of the system is required for self-organization of regulated structure, internal relative evaluation is performed. In the same way if we consider the variation in strategic diversity as a hypothetically generated process of self-organization, the self reference coupling structure is also a chained hypothetical process and any variation in the strategic diversity becomes an element of the intelligence system.

Sannomiya has tested the design and operation of system to flexibly respond to environmental change and discussed the simultaneous existence of autonomy and self in responding to environment change. He has used 2 group behavior models for explaining the emergence of cooperative behavior, but has assigned diversity an initial setting value for information exchange or awareness deciding mechanism. In response to this, we argue that diversity is an element that comes from within as a chain hypothesis generated process and corresponds to the system expansion and contraction. Based on the above concept, we verify the validity of the 3 design guidelines in system design for dynamic environment in which AGV runs and test the relationship of this concept to the actual engineering system by means of simulation.

Although it is necessary to identify the optimal parameters for design of distributed autonomous system, the identification method or reasons for suitability have not been discussed. In this research we propose and verify the method of identification of the optimum learning parameters for agent strategy reorganization for dynamic environment.

2. System Definition

2.1 Outline of the model

Let us consider an AGV transportation system as shown in Fig.2. In the model shown in Fig.2 (a), there are 20 nos. of AGV. The system consists of one Loading Station (hereafter LS) and one Unloading Station (US). Number of AGVs used for transportation remains constant.

The transportation rule is explained as follows: Each AGV is assigned a job in which one load is to be transported from LS to US in a repeatable cycle. 6 number types are used for load assignment, and AGV will unload to the US according to the required conveyance order of finite length. The system's purpose is to check that the success probability of Unloading by AGV increases as the conveyance order from US is satisfied. Every unloading count k_c set on the AGV will be counted as success ($r_c(k_c) = 1$) when the load type no. on the AGV matches with the load type no. requested by the US, and con-

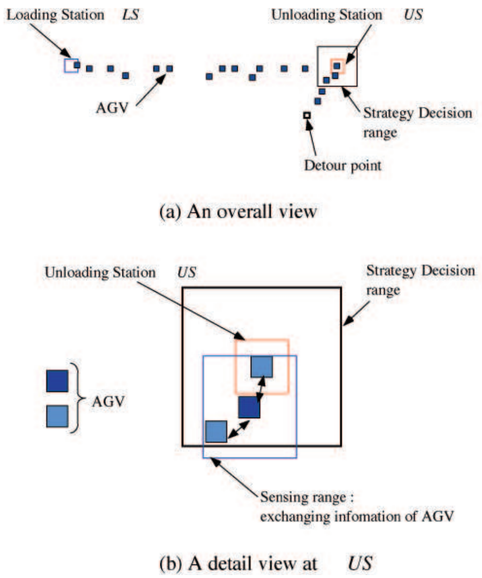

Fig. 2 System configuration of AGV transportation system

veyance volume will be counted up. If the load type on the AGV does not match with load type requested by the US, unloading count will be counted as failure ($r_c(k_c) = 0$) , and conveyance volume will not be counted up. The AGV can unload to the US irrespective of whether it gets an OK or NG reply, however, in case of OK reply, the output will be counted up and the load can remain on the US. In the case of NG reply, the load is immediately ejected from the US.

Loading scheme of AGV is explained as follows: The LS will supply the load requested by US in random sequence. AGV cannot select a desired load type no. from the LS. When the AGV goes to pick up the load, it will detect the load type number for the first time. Also the AGV cannot know the requested conveyance order from the US. Only when the AGV enters into the Strategy decision range as shown in Fig.2(b), will it be able to detect the previously OK load no. on the US.

Mutual interaction between AGV is explained as follows: As shown in Fig.2 (b), when an AGV enters the Strategy decision range, it can mutually interact with other AGV's in the sensing range. Through the information obtained in this interaction, an AGV can obtain its own strategy or internal status according to its own load number. The strategy involved here is to infer the next load to be unloaded to the US as a probability series to achieve the systems objectives. The AGV however cannot observe the load no. carried by other AGV or the reply received for an unloading operation $r_c(k_c)$ by other AGV. This includes the uncertainty involved in mutual interaction. It is possible to verify if strategy can be reorganized to achieve the system objectives in response to the uncertainty through this system. Based on the strategy of other AGV and own strategy obtained through mutual interaction, the AGV will decide if the load it is carrying can be unloaded to the US after the current unload operation is completed according to probability and will execute the unloading operation. If it cannot unload, it will make a detour to the Detour Point as shown in Fig.2 (a) and repeat the unloading operation to the US. After completing unloading, the AGV will move to the LS to pick up the load and repeat this cycle.

2. 2 Matrix representation of strategy

AGV Strategy decision method is described below. In this model the transport form is expressed as the unloading sequence based on the load type number. In formulating the strategy for AGV c, we let b_i be the load of Type i on AGV c, and w_j be the load of Type j on the US for N_b representing all loads and define estimated probability $p_{ij}(k_c)$ of unloading count k_c from b_i to w_j containing matrix $P_c(k_c)$. As the value of k_c is set individually for each AGV the value is not uniform for all AGVs. Using the estimated probability matrix, $P_c(k_c)$, if we take the decimal value of the closed interval$[0, 1]$, it is expressed as

$$P_c(k_c) = \left\{ \ p_{ij}(k_c) \ \right\} \in h^{N_b \times N_b}, \tag{1}$$

In our example, we set $N_b = 6$.

By using this estimated probability matrix $P_c(k_c)$, it can be expressed in probability terms as to which load should be unloaded next on the US. $P_c(k_c)$ is the information held independently by each AGV. $P_c(k_c)$ changes with the change in the unloading count k_c through the process of learning and interaction with other AGV. The initial value of estimated probability matrix when all the transport load numbers are N_b is:

$$P_c(k_c = 0) = \left\{ \ \frac{1}{N_b} \ \right\} \tag{2}$$

2. 3 Definition of internal status

The internal status on which basis it is possible to do the relative evaluation of the system purpose for each agent from the strategy $P_c(k_c)$ is described below: Internal state of the AGV is described as $\phi_c(k_c)$. The internal state $\phi_c(k_c)$ is defined as the degree of contribution to the system's objective. In this model, the AGV compares its own agent with that of other agent regarding its internal status $\phi_c(k_c)$. In this model, system's objectives are stipulated and this will not change due to change in environment. For this reason, it is not problematic to integrate the evaluation function that represents the degree of contribution to the systems objective. The internal state of an AGV c for an unloading count k_c is composed of $H_c(i, k_c)$ and

$S_c(i,k_c)$. $H_c(i,k_c)$ is the entropy of the estimated probability matrix $P_c(k_c)$ for load i on AGV and $S_c(i,k_c)$ is the successful unloading probability for load i on AGV. As the total number of loads is N_b, the entropy $H_c(i,k_c)$ is

$$H_c(i,k_c) = -\sum_{j=1}^{N_b} p_{ij}(k_c)\log_2 p_{ij}(k_c) \tag{3}$$

The successful unloading probability $S_c(i,k_c)$ is defined as the probability at which AGV c unloads b_i correctly according to the request from the US during r times.

$$S_c(i,k_c) = \sum_{\tau=k_c-r}^{k_c} f_c(i,\tau)/r \quad (k_c > r), \tag{4}$$

where

$$f_c(i,k_c) = \begin{cases} 1 & iff\ AGVc\ succeeds(r_c(k_c)=1) \\ 0 & iff\ AGVc\ fails(r_c(k_c)=0) \end{cases} \tag{5}$$

Hence internal evaluation q_i of AGV c in response to strategy i is:

$$q_i(k_c) = (1-\alpha H_c(i,k_c))S_c(i,k_c). \tag{6}$$

The internal state vector $\phi_c(k_c)$ considered as an element q_i is assigned as follows:

$$\phi_c(k_c) = [q_1(k_c)\ldots q_i(k_c)\ldots q_{N_b}(k_c)]^T. \tag{7}$$

where α is the normalization factor of the entropy expressed by the following equation:

$$\alpha = \frac{1}{-\sum_{i=1}^{N_b}\frac{1}{N_b}\log_2\frac{1}{N_b}} = \frac{1}{\log_2 N_b}. \tag{8}$$

$\phi_c(k_c)$ is an important element for AGV interaction. When the internal state is $\phi_c(k_c)$ and internal evaluation $q_c(k_c)$ is high, then entropy of estimated probability matrix $P_c(k_c)$ is low. This indicates progress of the strategy organization and an increase in successful unloading probability $S_c(i,k_c)$. System objectives are met when unloading is done according to the conveyance order. By using this internal state, AGV can observe the contribution degree of the system autonomously without using global evaluation function.

2.4 Learning scheme of the strategy

The method of learning by each AGV is described below. After unloading, the AGV learns through its own behavior to realize self organizing of the AGV transport form. The focus of learning is the strategy $P_c(k_c)$. In case AGV does not unload, as it just moves to a Detour Point, there will be no learning. After a load w_j of type j on US, AGV unloads load b_i of type i on the US, and reply $r_c(k_c)$ will be sent from the US to the AGV. From this, the AGV can by itself observe the success or failure of unloading. AGV learning will occur based on the value of the reply $r_c(k_c)$. The element $p_{ij}(k_c)$ of the estimated probability distribution $P_c(k_c)$ is updated using the value derived in Eq.(7), learning parameters during unloading success or failure and the internal status.

i) If unloading is successful ($r_c(k_c)=1$):

$$p_{ij}(k_c+1) = p_{ij}(k_c) + \gamma_{OK}(1-p_{ij}(k_c))(1-q_i(k_c)) \tag{9}$$

$$p_{uj}(k_c+1) = p_{uj}(k_c) - \gamma_{OK}p_{uj}(k_c)(1-q_u(k_c)) \quad (u\neq i,\ u=1,2,\ldots,N_b) \tag{10}$$

ii) If unloading is a failure ($r_c(k_c)=0$):

$$p_{ij}(k_c+1) = p_{ij}(k_c) - \gamma_{NG}p_{ij}(k_c)q_i(k_c) \tag{11}$$

$$p_{uj}(k_c+1) = p_{uj}(k_c) + \gamma_{NG}(1-p_{uj}(k_c))q_u(k_c) \quad (u\neq i,\ u=1,2,\ldots,N_b) \tag{12}$$

As described above, in this system in order to create a strategy for the dynamic environment based on L_{R-P} learning [7], internal status $\phi_c(k_c)$ (contribution to system through self organizing) is used as the basis for learning .

2.5 Interaction scheme of the agents

Interaction process is described below. The information exchanged mutually between the AGVs is $\phi_c(k_c)$ and $P_c(k_c)$. AGV obtains the estimated probability matrix $p_d(k_c)$ and internal status $\phi_d(k_c)$ (on condition that $c\neq d$) from the m number of AGV in the sensing range as shown in Fig.2 (b). In this research in order to realize dynamic reconfiguration of the conveyance form, it is necessary to stipulate the group behavior of the AGVs. For this, we do not think it is appropriate to use an absolute evaluation based on restrictive conditions of a large area. That is to say, when there is no objective valuation standard, the value evaluation will be relative and susceptible to fluctuate. For this reason, the pattern of interaction between the AGV is as described below and set such that AGV c strategy $P_c(k_c)$ is propagated to other AGV as well.

$$P_c(k_c+1) = P_c(k_c) + \xi(\phi_c(k_c) - \bar{\phi}_d(k_c))(P_c(k_c) - \bar{P}_d(k_c)) \tag{13}$$

$$\bar{P}_d(k_c) = \frac{\sum_{d=1}^{m} P_d(k_c)}{m}, \qquad \bar{\phi}_d(k_c) = \frac{\sum_{d=1}^{m}\phi_d(k_c)}{m} \tag{14}$$

where, ξ is a parameter for change in interaction. ($\xi > 0$)

From equation (13) the interaction that takes place in this model consists of m number of AGV interacting with each other. Through the dynamics of the system, the strategy of the AGV whose contribution to the system is high is propagated to other AGV. This allows each AGV to increase its level of contribution to the system by itself.

2.6 Expression of unloading selecting probability

From the above, AGV will execute the below unloading from the derived strategy $P_c(k_c)$. When the AGV enters the Strategy decision range, US transport load information $L_{US}(k_c)$ can be obtained from the US. $L_{US}(k_c)$ is the binary encoded information of element N_b. For eg. if the load no. on the US is $j = 1$. $L_{US}(k_c \mid j = 1)$ will equal to row vector $[1\ 0\ 0\ 0\ 0\ 0]^T$. Based on this information, AGV c will derive the selection probability vector $Y_c(k_c)$ from the strategy $P_c(k_c)$. P_c is the strategy for determining when b_i must unload after w_j, so by exploiting L_{US} in P_c, the selection probability vector for b_i corresponding to w_j can be obtained.

$$Y_c(k_c) = \epsilon P_c(k_c) L_{US}(k_c) \tag{15}$$

On condition that ϵ is the norm standard coefficient of vectors.

The selection probability vector $Y_c(k_c)$ is loaded with load information on the AGV $L_{AGV}(k_c)$ to derive the selection probability $y_{cij}(k_c)$.

$$y_{cij}(k_c) = (Y_c(k_c), L_{AGV}(k_c)) \tag{16}$$

where $L_{AGV}(k_c)$ like $L_{US}(k_c)$ is the binary coded information of the element N_b.

AGV executes unloading according to the selection probability $y_{cij}(k_c)$.

3. Condition of the Simulation

3.1 Definition of transportation order

The requested conveyance order for the transport load is expressed by the array O of several records N_o. In this simulation, $N_o = 10$ and $O = \{1\ 1\ 2\ 3\ 4\ 2\ 6\ 3\ 5\ 6\}$. When the load type reaches the end of the array, the next load type will be at the top of the array. Fig.3 expresses the requested conveyance order and the estimated probability matrix P^* for the exercise used in this research. P^* in Fig.3 indicates the number corresponding to the load b_i on the AGV and the load w_j on the US after which if unloading is done to the US, the probability of successful response would be high. This is displayed as a bar graph based on the distribution status derived from the requested conveyance order O. The designed successful unloading probability of the whole system $\hat{E}$ is expressed by using the designed probability matrix P^* as follows: We express $p(b_i)$ as an appearance probability of the load b_i on the AGV, p^*_{ij} is the estimated probability from w_j to b_i and collection of load w_j on the US during successful unloading of b_i by AGV as

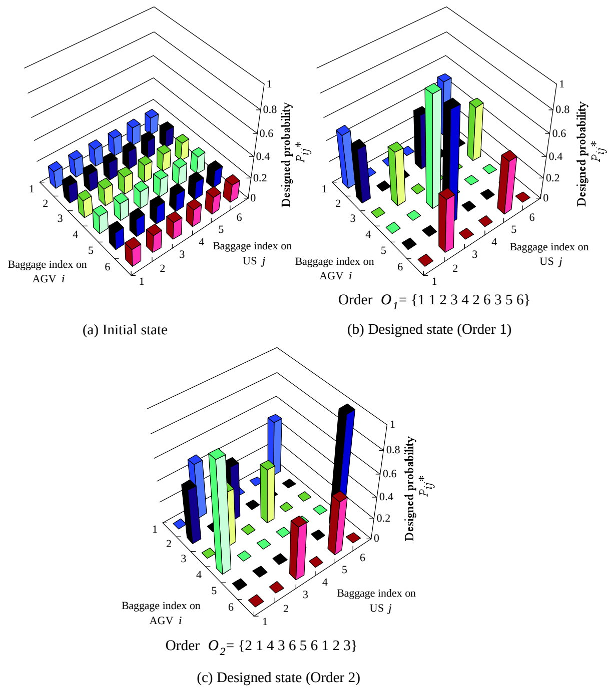

Fig. 3 Desired probability matrix of experiment

$V_i = \{j = w_j | b_i \rightarrow w_j\ permitted\}$. The expected successful unloading probability $\hat{E}$ for the entire system based on the requested conveyance order is expressed as follows:

$$\hat{E} = \sum_{i=1}^{N_b} \sum_{j \in V_i} p(b_i) p^*_{ij}, \tag{17}$$

where N_b is the transport load type within the conveyance order.

From the designed probability matrix P^* in Fig.3, $N_b = 6$, if load having numbers 1, 2, 3, 6 appear twice, the appearance probability is $\frac{1}{5}$, and as load numbers 4, 5 appear only once, the appearance probability is $\frac{1}{10}$. By using these parameter values, it is possible to derive the designed successful unloading probability $\hat{E} = 6$. In this simulation, as the requested conveyance order has some identical numbers, the conveyance order organized by the AGV will be complex.

3.2 Definition of dynamical environment

The setting of the dynamic environment for this research is described below. The change in environment is done within a range not exceeding the degree of freedom of the agent (strategy space). For eg. the load type will not be increased beyond 6 and conveyance order monitored. In this system, as shown in Fig.3, the requested conveyance order on the US is taken as $\{1\ 1\ 2\ 3\ 4\ 2\ 6\ 3\ 5\ 6\}$, and the requested conveyance order from the LS to the AGV is set at random, to supply the load and form the conveyance order. After this at step 100000, the the requested conveyance order is modified to $\{2\ 1\ 4\ 3\ 6\ 5\ 6\ 1$

2 3}, and it is verified if the AGV can restructure the conveyance order. The AGV makes one movement at each step or executes one grid movement. Since there are 50 grids between the LS and the US, AGV movement will be complete in about 50 steps. This simulation was done up to 350000 step.

The following items are considered for verification in the defined dynamic environment.

(1) Verification of dynamic recompilation of strategic diversity and variation of the conveyance order.

(2) Verification of the time evolution of the variation process in strategic diversification.

4. Simulation result

4.1 Expansion-contraction of strategic diversity adapting to dynamic environment

Simulation result based on the environment setting as described in Section 3.2 is displayed in Fig.4. Fig.4 is a graph displaying the strategic diversity and the probability of unloading success of AGV within a dynamic environment where the requested conveyance order is modified. The value of the learning parameters γ_{OK} and γ_{NG} is set at 0.2 and 0.15 respectively. This graph indicates that through the process of self destruction of strategy, the AGV reorganizes a strategy suitable to the dynamic environment. By this the unloading success probability is restored and collective autonomous behavior is obtained. If we focus on the unloading success probability graph shown in Fig.4, the recovery of the unloading success probability indicates that the collective autonomous behavior obtained through the process of learning by each agent and mutual interaction, meets the system's objectives. However, the learning algorithm is programmed such that unloading success probability increases for each agent. Merely by looking at the unloading success probability graph, it is not possible to judge that self-organization has been achieved from the collective autonomous behavior for the system as a whole from the recovery of the unloading success probability. It is thus possible to conclude that the recovery of the unloading success probability is expressed as self organization of the system and self organization index is required in order to examine the behavior of the entire system from the point of view of the design method of the Distributed autonomous system.

The strategy of agent c is defined as the selection probability matrix P_c input vector as x and output vector as y and assuming there is a static linear relationship between the input and output, the agent behavior can be expressed as follows.

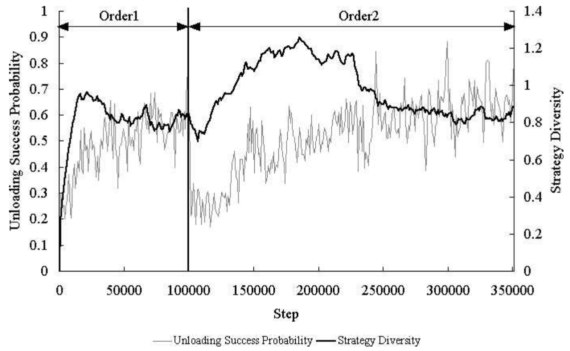

Fig. 4 Time evolution of Strategy diversity and Unloading success probability ($\gamma_{OK} = 0.2$, $\gamma_{NG} = 0.15$)

$$y = P_c x \tag{18}$$

From the fact that all the agents in this system are set with a simple job objective possessing identical evaluation standards, it is possible to say that agent strategy will converge to a specific matrix. By deriving the difference in the selection probability matrix of each agent, the average value for one AGV can be calculated from the total and defined as the strategy diversification for that AGV. This is expressed in Equation (19).

For a system with n agents:

$$\delta = \frac{\sum_{d}^{n}\sum_{c\neq d}^{n}\|P_c - P_d\|}{n(n-1)}. \tag{19}$$

Where Frobenius- norm is used for the matrix norm and expressed as follows:

For matrix $P = \{p_{ij}\}(\in N \times N)$:

$$\|P\| = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}|p_{ij}|^2}. \tag{20}$$

We consider the strategy diversity δ derived from equation (19) as the index of self-organizing and proceed with the below study. Strategy diversity in Fig.4 expresses the time evolution δ. In Fig.4, the value of δ at Step 0 is 0 and indicates that all the agents are having the same strategy. From the time evolution graph δ shown in Fig.4, by expanding the autonomous strategic diversity from the uniform strategic pattern, the strategy search range is widened and self-organization of strategy formation occurs. Even if there is a change in the environment over 100000 steps that affects the strategy, and although this is not informed to other AGV, through the process of autonomous contraction and expansion of strategy, each AGV will destroy the old strategy and reformulate a new one. This can be described as a phenomenon of self-organization that has been achieved through the

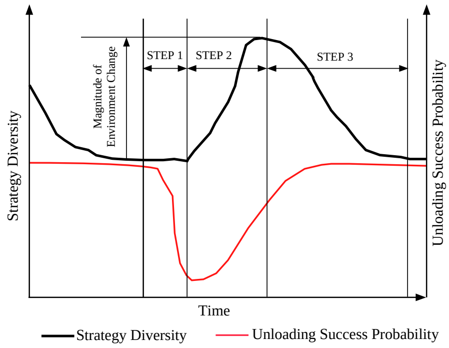

Fig. 5 Schema of diversity-regulation process

algorithm proposed in this research. It is necessary to know the process of self destruction and reorganization of strategy [2] held by each agent in order to verify the suitability to the dynamic environment for the intelligence system, but in order to consider within the framework of autonomous group behavior, it is necessary to explain the expansion and contraction process for strategic diversity occurring during each agent behavior from the self destruction and reorganization process of strategy [9]. By clarifying the mechanism suitable for the dynamic environment from the process, it is possible to design and generalize a distributed autonomous system for the dynamic environment and there would be no need for creating a numerical model of the change in environment to provide an example.

The expansion and contraction process of strategy diversification corresponding to the change in environment is schematically expressed in Fig.5 as follows.

As shown in Fig.5, the expansion and contraction process of strategy diversification can be bifurcated into three levels. In the first level, each agent exhibits behavior holding the strategy prior to environment change, so the strategic diversity exhibits a constant value with a decreasing unloading success probability. From this it is possible to understand that each agent does not recognize the change in environment at the first level. At the second level, the fact that the strategic diversity increases indicates that each agent recognizes the environmental change and expands the strategic space for creating a new strategy. This leads to the emergence of an individualized entity for sure acquisition of strategy matching with the system objectives, and also leads to an increased unloading success probability. The strategic diversity can be expanded up to a certain value according to the change in environment based on the learning co-efficient and mutual interaction. The maximum value of strategy diversification will be the maximum change in the environment for the system. The size of change in environment will vary according to the conditions for dynamic environment (change in conveyance order), mutual interaction between the agents, and density of presence probability, strategy searching space etc. Each element of the system in intertwined in a non-linear fashion and has a big effect on the time needed for strategy reformulation. At the third level, the number of AGV with unloading success probability satisfying or meeting the system's objectives increases and concentrated at the value where desired probability rate is achieved. Even at this time, there is some dispersion in the strategy values and process of expansion and contraction of strategy still continues. Strategy diversification continues to expand and contract as strategy is reformulated to meet system objectives and when strategy converges to a single value, for the first time, it could be said that strategy reformulation is complete.

As can be seen in Fig.4, the convergence value for δ is 0.8 and we come to know that strategy is not converging to 0. This means that in this research model, although all agents are working towards a single objective, they do not move to a specific distribution range. However, from the fact that unloading success probability is converging to about 0.6, and from the fact that δ indicates a stable value of 0.8, the agents distribution strategy comprises multiple formats and is a stable value. The strategy distribution status for system organization is diversified and suggests that labor differentiation occurs over the entire system strategy.

In order to know if the optimum value for strategy diversification is displayed for this system, it is necessary to judge if each agent within the system has acquired behavior representing global solutions. The discussion may differ depending on what can be categorized as global and what can be categorized as local, but assumption of a stable global solution within a dynamic environment is complicated and its existence is not clear. For this reason, this research does not discuss the optimization of collective autonomous behavior. This research suggests that through the process of expansion and contraction of the strategic diversity process, a dynamic coherence is created within the changing environment that is the basis of the self referential structure of the intelligence system. From this it is possible to discuss if the optimization of the evaluation of the autonomous behavior due to a drastic change in the environment for the entire system is appropriate or not.

4.2 Identification of parameters corresponding to dynamic environment

It is necessary to identify the learning parameters for de-

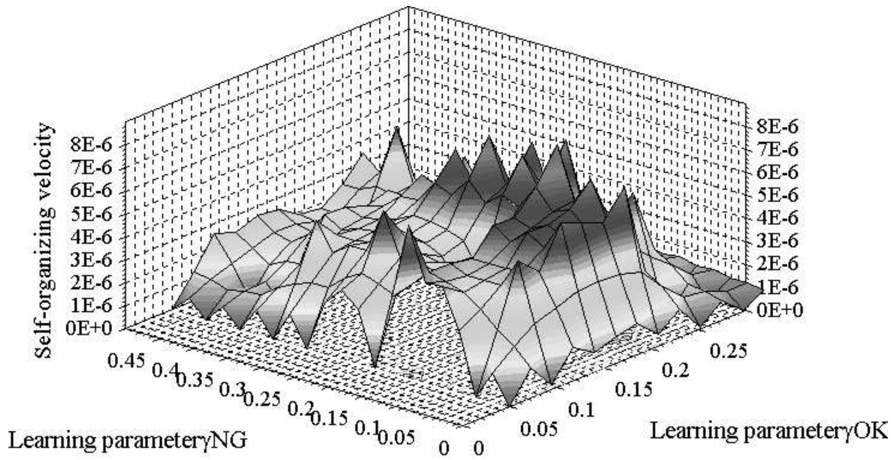

Fig. 6 Self-organizing velocity distribution based on learning parameter (γ_{OK} and γ_{NG})

sign of an distributed autonomous system for the dynamic environment. However deriving the operator of the entire system from the behavior of each agent that is the mutual interaction between the agents, density of existence probability, strategy searching space etc. i.e. to interlink the operator of each system in a non-linear fashion is presently complicated. If we measure the slant of the line from the maximum value of the expansion and contraction speed of strategy diversification to the intersection between the convergence value and decreasing bending line as shown in Fig.4, the speed of expansion/contraction in a static environment is 8.72×10^{-6}, and 8.23×10^{-6}, when there is a change in the environment indicating that there is no difference in the expansion and contraction speed due to a change in the environment. We think that the contraction and expansion of strategy diversification would depend on the speed of self-organization existing in the dynamics of the strategy searching space for each agent, learning speed or mutual interaction between the agents and the system size. The speed of self-organization is reflected when the strategy formulation speed by each agent is more than the environmental noise generated from the mutual interaction between agents, environmental uncertainty, and massive strategy searching space. The expansion and contraction speed of strategy as shown below is called as the (Self-organizing velocity). In this system, the learning parameters values γ_{OK} and γ_{NG} have been revised respectively from 0.05 to 03 and from 0 to 0.5, and the distribution range of the self-organization velocity was compared. This is shown in Fig.6. The self-organizing velocity is the change in volume of strategic diversity δ at each step for time averaged from the maximum value of strategic diversity to the convergence value. According to Fig.6, the

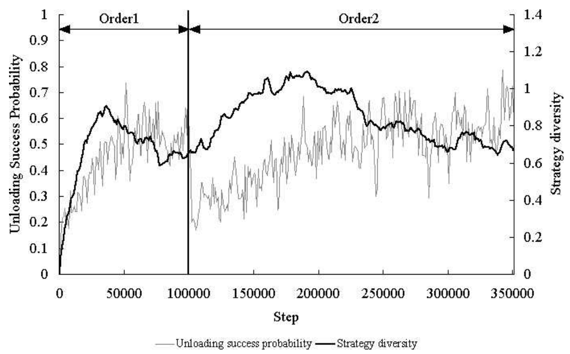

Fig. 7 Time evolution of Strategy diversity and Unloading success probability ($\gamma_{OK} = 0.1$, $\gamma_{NG} = 0.1$)

value of the self organizing velocity is 8.72×10^{-6} when the $\gamma_{OK} = 0.2$, and $\gamma_{NG} = 0.15$ and this is the maximum. Fig.4 shows the strategy diversity and unloading success probability by using these parameters and this indicates that the system behavior is in the most suitable dynamic environment in the area range set in fig.6. Fig.7 shows the simulation result when $\gamma_{OK} = 0.1$ and $\gamma_{NG} = 0.1$. The self organizing velocity in Fig.7 was $3.90 \times$ 10-6. As shown in Fig.7, it takes approximately 8000 steps from the time the system starts operating till the convergence of strategy diversification. If we compare this with fig.4 with self organizing velocity of $\gamma_{OK} = 0.2$, and $\gamma_{NG} = 0.15$, the difference for reaching the convergence level is only about 1000 steps. However, if we observe the system behavior when a change in environment is made after the system is in operation for more than 100000 steps we can see that it takes about 200000 steps for the strategy diversification to once again reach the convergence level. In Fig.4 it takes only 150000 steps for reaching the strategy differentiation to reach the convergence level creating a big difference in dynamic reformulation.

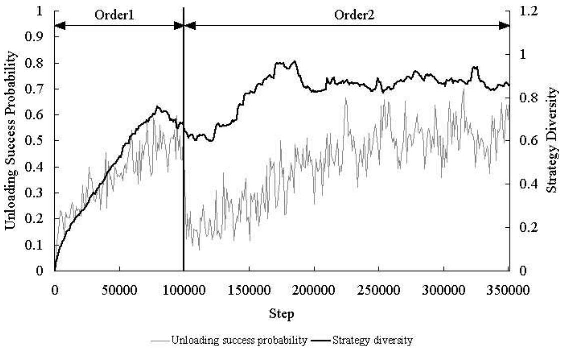

Fig. 8 Time evolution of Strategy diversity and Unloading success probability ($\gamma_{OK} = 0.05$, $\gamma_{NG} = 0.05$)

A big delay is generated in the unloading success probability similar to the transition in strategy diversification. Also from Fig.7 it is possible to understand that time is required in the process for increasing the strategy diversification. In the same way, using the learning parameter representing a small value of the self-organization velocity does not give the desired results in terms of unloading success probability to for achieving system objectives or strategy diversification as the index of organization. The system behavior when $\gamma_{OK} = 0.05$ and $\gamma_{NG} = 0.05$ is represented in Fig.8 for reference. At this time, the self-organizing velocity from Fig.6 is 0 (There is no convergence of strategy diversification within 80000 steps). From Fig.8 it is possible to confirm that there is practically no expansion or contraction of strategy diversification occurring the dynamic environment and no strategy reformulation for recovering the unloading success probability.

From the above, we can say that it is possible to define the self-organizing velocity and derive the features of distribution of the self organizing velocity in a static environment from which it would be possible to derive the learning parameters for generating a high self organizing velocity in the static environment. By the derivation of these learning parameters, it is possible to define the most suitable parameters for realization of appropriate behavior in the dynamic environment for design of autonomous distribution system. Based on the learning parameters or other restrictive conditions, and deriving the distribution status of the self organizing velocity, it would be possible to predict how the system behavior changes according to the parameters in the dynamic environment. This has been verified in the simulation results.

5. Conclusion

In this research, we have defined strategic diversity as the difference in strategy between all the agents and by verifying the time evolution of strategic diversity occurring in the dynamic environment, it is possible to show that organization takes place through the collective autonomous behavior that meets system objectives according to expansion and contraction of strategy diversification. Definition of system parameters is one of the most important items in system design. Also this task is enormous in the case of large-scale system. In this research we have defined the expansion and contraction speed of strategic diversity as self organizing velocity and proposed a method for defining the most suitable parameters derived from the distribution range of self organizing velocity which in turn is derived from the definition of parameters for a static environment. Through this method, simulation results have shown that the highest dynamic reorganization of strategy can be achieved by using the highest self-organizing velocity.

However, if the self-organizing velocity is too high, the strategy may collapse. We therefore need to prove quantitatively the ideal conditions for defining the optimum parameters corresponding to the maximum self-organizing velocity in the dynamic environment. Also it is necessary to consider and analyze the maximum value of strategic diversity and change in environment and the correlation between the convergence value of the strategic diversification and restrictive conditions of the system like learning parameters.

In order to design a self creating model for facilitating autonomous adjustment of parameters in a dynamic environment, analysis of a numerical model of time evolution of operator theory in function spaces. This is a topic that needs evolution of a different topic.

References

1) Higashi, T., Sekiyama, K., and Fukuda, T. Autonomous strategy organization in the dynamical quantitative environment. In *Proceedings of the 1998 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*, pages 1547–1552, 1998.
2) Higashi, T., Sekiyama, K., and Fukuda, T. An autonomous strategy reconfigurable agv transportation system using self-referential interaction (in japanese). *Transactions of The Japan Society of Mechanical Engineering (C series)*, 65(638):4121–4128, 1999.
3) Higashi, T., Sekiyama, K., and Fukuda, T. Self-organizing control of carrier se-quence in agv transportation system. In *2000 IEEE International Conference on Industrial Electronics, Control and Instrumentation (IECON-2000)*, pages 706–711, 2000.
4) K.Hirasawa, , J.Hu, , J.Misawa, , J.Murata, , and H.Katagiri, . Rasid training of multi agent systems with fuzzy inference-based mutual interactions. *Machine Intelligence and Robotic Control*, 2(1):17–25, 2000.
5) Kimura, H. and Kobayashi, S. Reinforcement learning for continuous action using stochastic gradient ascent,. In

Proc. of 5th Int. Conf. on Intelligent Autonomous Systems, pages 288–295. IEEE, January 1998.

6) Kurabayashi, D., Ota, J., Arai, T., and Noguchi, K. Behavior of a group composed of robots with heterogeneous motion agorithms. In *Proc. IEEE/RSJ Int. Conf. on Inteigent Robots and Systems*, pages 241–246, 1999.
7) Narendra, K. and Thatahchar, M. *Learning Automata an Introduction*. Prentice-Hall, 1989.
8) Sannomiya, N., Tian, jie Y., and Nakamine, H. Emergence of cooperation in fish behavior models (in japanese). *Transactions of the Society of Instrument and Control Engineers*, 35(11):1370–1376, 1999.
9) Sekiyama, K. Self-referential dynamics with situated observation in collective agent systems. In *1999 IEEE International Conference on Systems, Man, and Cybernetics*, pages I–244–I–249, 1999.
10) Watanabe, M. *Studies on AGV System with Autonomous Agent*. PhD thesis, Hokkaido University, 2003.
11) Yoshii, S., Ohashi, S., and Kakazu, Y. Modeling of emergent ecology for simulating adaptive behavior of universal computer programs. In *From Animals to Animats 5, Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*, pages 303–308, 1998.

Toshimitsu Higashi (Member)

received the Master of Engineering degree from Waseda University, Tokyo, Japan in 1992. He joined Murata Machinery, LTD., Aichi, Japan in 1992. Meanwhile, he received the Doctor of Engineering degree from Nagoya University in 2000. At present, he is in Logistics and Automation Division in Murata Machinery, LTD., Aichi, Japan, mainly engaging in the research fields of analysis and design of logistics systems, self-organizing manufacturing system and collective autonomy of distributed autonomous robotic system.

Kosuke Sekiyama (Member)

Kosuke Sekiyama received the B.E. and M.E. degrees both in mechanical engineering in 1992 and 1994 from Nagoya University, Nagoya, Japan respectively. He received his Ph.D. degree in engineering from Nagoya University in 1997. He was a Research Associate in the Department of Micro System Engineering, Nagoya University from 1997 to 1998. He joined the Science University of Tokyo, Suwa College in 1998. He is currently an Associate Professor in the Department of Human and Artificial Intelligent Systems, Fukui University, Fukui, Japan. His research interests include distributed autonomous systems and emergent systems.

Toshio Fukuda (Member)

Prof. Toshio Fukuda graduated from Waseda University, Tokyo, Japan in 1971 and received the Master of Engineering degree and the Doctor of Engineering degree both from the University of Tokyo, in 1973 and 1977, respectively. At present, he is Professor of Department of MicroSystem Engineering, Nagoya University, Japan. He is mainly engaging in the research fields of intelligent robotic system, self-organizing system, micro robotics, robotic system under hostile environment, bio–robotic system, neuromorphic intelligent control, fuzzy control, control of mechanical systems and technical diagnosis. He is a member of SICE, RSJ, JSME, IEEE, a Fellow of SICE, IEEE, and currently IEEE Nanotechnology Council President.

Reprinted from Trans. of the SICE, Vol.37, No.5, 438/445, (2001)