Evaluation of an economic model composed of producer agents

Nao ITOU, Yoshinobu MAEDA and Toyohiko HAYASHI

Guraduate School of Science and Technology, Niigata University

8050 Ikarashi-2, Nishi-ku, Niigata 950-2181, Japan

email:nao@gis.ie.niigata-u.ac.jp

Abstract—**In recent years, the existence of a "social divide", comprising factors such as income and professional status, has been noted as one significant type of social issue. It has been stated that, among the disparate groups, a member belonging to one group cannot move to any other groups. Such a rigidity in terms of social status results in non-activation of the economy. In this study, we have suggested a dynamic economic model described by a multi-agent system, and have evaluated its dynamics in order to try to understand the mechanisms by which how the social divide emerges within the model. We used Gini's coefficient to evaluate the social divide and its economic efficiency. As a result, it is suggested that economies under conditions of low competitiveness, being a state composed of relatively more consumer agents than producer agents, display a higher negative relationship between Gini's coefficient and economic efficiency than those under conditions of high competitiveness.**

I. INTRODUCTION

In recent years, the existence of a "social divide", comprising factors such as income and professional status, has been noted as an emerging social issue. This "divide" is defined as a difference in price, license, classification, or level and amount of information within the same class. When the income gap within a society is large, i.e., the society is in a state where 'blue-collar' members are not able to escape from the lowest levels in the hierarchy, then the motivation for social development will atrophy and the society will become economically dormant. In order to develop and maintain an active society, both ultra-strong and ultra-weak levels of divide are undesirable. So, is it possible to maintain an appropriate level of divide in order to encourage social development?

There are various theories as to the cause of the emergence of a social divide, such as academic background, aging of society and part-time employment. However, the mechanisms for the emergence of the social divide that are caused by the economic system are unclear. In general, a system has a tendency to generate a divide to adjust its overall efficiency. For example, multi-cellular organisms differentiate their cells into many functionally different parts to maintain the individual's life. Similarly, an economic system is composed of characteristically different individuals.

The main purpose of this study is to investigate the relationship between the income divide and the economic efficiency in an economic model, and is presented in this paper using multi-agent simulation (MAS). MAS is often used when studying the dynamic features of social phenomena [1], [2].

In our economic model, the agents learn their action using reinforcement learning in order that they may be adapted to the environment.

II. MULTI-AGENT SYSTEM

A multi-agent system (MAS) is composed of many agents, and achieves a task that is difficult to achieve for a single agent or a monolithic system. In a MAS, an artificial society comprising many autonomous behavioral agents interact with each other in the simulation [1], [2]. We can contrast the phenomena simulated by a MAS with those of the real world. However, it is difficult to identify all of the relevant states because the environment in the MAS changes depending on the agent's global features [3]. In this study, we constrain the agent to acquire an action rule, adapting it to the environment by using reinforcement learning.

III. REINFORECEMENT LEARNING

A. Definition

Reinforcement learning is a type of machine-learning that treats the problem where an agent in an environment recognizes the current state, $s_t \in S$, and decides the action, $a_t \in A$, to be taken [4]. When the agent chooses the action $a_t \in A$, the state of the environment translates to $s_{t+1} \in S$ and the agent receives a reward (Fig.1). Reinforcement learning makes the agent learn the policy that offers it the maximum reward through its actions. Here the environment is formulated as a Markov decision process that has a finite number of states. Reinforcement learning is thought of as a type of 'dynamic programming'. It is different from supervised learning, because such learning is not derived from a teacher. It has an important feature in that the agent can choose an action by developing an unknown-learning-area and an action using a known-learning-area in a balanced manner. Therefore, reinforcement learning is often used to achieve an action rules for a robot under an unknown environment. Thus, many methods of reinforcement learning have been proposed [5], [6], [7], [8], [9], [10], [11], especially in the following section, where profit sharing achieves a higher performance than alternatives when reinforcement learning is applied to the MAS' environment [12], [13]. Therefore, we have used profit sharing in this study.
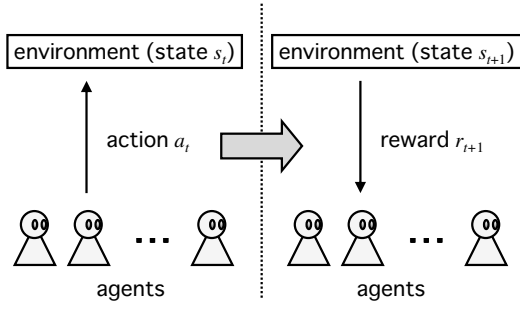
Fig. 1. Transition of a state in reinforcement learning.

Fig. 2. Flow of merchandise in a producer economic model.

B. Profit Sharing

Profit sharing is one type of method for reinforcement learning [12], [13]. In profit sharing, the agents are assigned a Q-value which shows the efficacy of a rule that is based on the agent's actions, and the Q-value is updated after obtaining the reward. The rule is a pairing of the state $s_t \in S$ and the action $a \in A$, which the agent can choose. With a state $s_t \in S$ assigned to the environment, the agent stochastically chooses the action $a \in A$, based on policy $\pi(s_t, a)$. We then apply the Boltzmann choice method [14].

$$\pi(s_t, a) = \frac{e^{Q(s_t, a)/T}}{\sum_{a \in A} e^{Q(s_t, a)/T}}, \tag{1}$$

where policy $\pi(s_t, a)$ is the probability of choosing action a in response to state s_t, and T is a temperature parameter. The agent implements action a, and, as a result, the agent receives the reward r_t. The Q-value, which takes the action a for state s_t, is updated as follows,

$$Q(s_t, a_t) \leftarrow (1 - \beta)Q(s_t, a_t) + \beta r_t, \tag{2}$$

where $\beta(0 \leq \beta \leq 1)$ represents the learning-rate.

Profit sharing is a type of non-boot-strap reinforcement learning, i.e., it does not require the next state's Q-value when updating. Therefore, it shows a higher performance than the boot-strap method when reinforcement learning is applied to an MAS environment.

IV. PRODUCER ECONOMIC MODEL

In the mechanism of the emergence of a social divide, it is foreseen that the mechanism is included in the economic system. In this study, in order to restage the emergence of social divides as macro phenomena that are caused by the action of micro economic entities, we have structured an economic model using a MAS. In the model, we set 'multi-merchandises' and 'multi-agents'. The merchandises are defined by the combination and permutation of two types of primordial merchandise. There are three types of agent; namely the supplier agent, the producer agent and the consumer agent. A supplier agent produces unlimited primordial merchandise, and sells it. A producer agent purchases merchandise comprising materials for his product, and he produces his product by using materials and then sells it at a price. The consumer
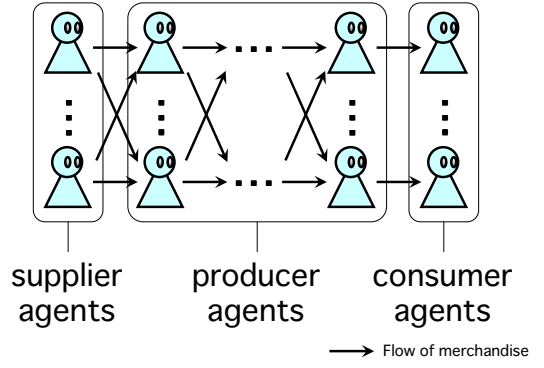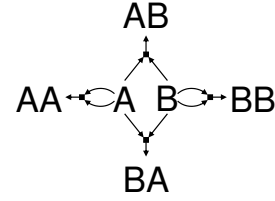

Fig. 3. Combination of materials and products.

agent purchases the merchandise that is finally produced by the producer agents. Fig.2 shows the flow of merchandise in a producer model. The producer agents learn appropriate action-rules concerning whether or not to purchase the material merchandise and how to price their product merchandise by reinforcement learning. The supplier agents and the consumer agents act on previously-determined action rules, i.e., not learned. In this producer economic model, we focus on the actions of the producer agents in the economic system. Here, we observe income division between the producer agents and the economic efficiency in this system, and evaluate the relationship between them.

A. Setting of merchandise

In this model, there are two types of primordial merchandise. We name these A and B. Merchandise AB is produced by combining A and B. In addition, merchandise BA is produced by combining A and B. Here, two kinds of merchandise can be produced by permutation. However, when the two materials are the same, such as both A and A or both B and B, only one kind of product merchandise can be produced, such as AA or BB (Fig.3). Here, we describe merchandise A and B as 'mono-bona merchandise', and merchandise AB, BA, AA and BB as 'di-bona merchandise'. TABLE 1 shows the product merchandise and the combinations of the materials that constitute product merchandise up to the 'tetra-bona level'.

B. Setting of agents

The supplier agents sell merchandise A or B at a pre-established price, p_A. The consumer agents purchase pre-

	Product	Material
Di-bona	AA	A+A
	AB	A+B
	BA	B+A
	BB	B+B
Tri-bona	AAA	A+AA,AA+A
	AAB	A+AB,AA+B
	ABA	A+BA,AB+A
	ABB	A+BB,AB+B
	BAA	B+AA,BA+A
	BAB	B+AB,BA+B
	BBA	B+BA,BB+A
	BBB	B+BB,BB+B
Tetra-bona	AAAA	A+AAA,AA+AA,AAA+A
	AAAB	A+AAB,AA+AB,AAA+B
	AABA	A+ABA,AA+BA,AAB+A
	AABB	A+ABB,AA+BB,AAB+B
	ABAA	A+BAA,AB+AA,ABA+A
	ABAB	A+BAB,AB+AB,ABA+A
	ABBA	A+BBA,AB+BA,ABB+A
	ABBB	A+BBB,AB+BB,ABB+B
	BAAA	B+AAA,BA+AA,BAA+A
	BAAB	B+AAB,BA+AB,BAA+B
	BABA	B+ABA,BA+BA,BAB+A
	BABB	B+ABB,BA+BB,BAB+B
	BBAA	B+BAA,BB+AA,BBA+A
	BBAB	B+BAB,BB+AB,BBA+B
	BBBA	B+BBA,BB+BA,BBB+A
	BBBB	B+BBB,BB+BB,BBB+B

established merchandise from the producer agent who sells it at the cheapest price.

The producer agent has the following properties,

M_1 : Material merchandise 1

M_2 : Material merchandise 2

P : Product merchandise $M_1 + M_2$.

v_{M_1} : Minimum price of material merchandise 1

v_{M_2} : Minimum price of material merchandise 2

v_P^{t-1} : Maximum price of product merchandise that was purchased on the last occasion.

s : $s = (v_{M_1} + v_{M_2}, v_P^{t-1})$. State that is recognized by the producer agent

a : $a = (Buy, v_P)$. Action that is chosen by producer agent

Buy : Variable that decides purchasing of material merchandise. It has values 0 or 1. When the agent does not purchase, Buy is 0. When the agent purchases, Buy is 1

v_P : Sale's price at next occasion. $0 \leq v_P \leq 30$.

C : Cost. If $Buy = 0$, $C = 0$ because the agent does not purchase anything. If $Buy = 1$, $C = v_{M_1} + v_{M_2}$.

I : Sale. If the agent's product merchandise is purchased, $I = v_P$, otherwise $I = 0$.

R : Income. $R = I - C$. It is also a reward on reinforcement learning.

$Q(s, a)$: Q-value of state s for action a.

C. Action of agents

Step1. Each property of the agent is initialized.

Step2. An agent is selected from the producer and the consumer agents.

Step3. The selected agent recognizes the state s. The agent implements Step4 and Step5 when this agent is the producer agent, M_1 and M_2 are sold. If this agent is the producer agent and M_1 and/or M_2 are not sold, he does not implement Step4 and Step5, because in this case he can't produce his product. If this agent is the consumer agent, he purchases pre-established merchandise from the producer agent who sells at the cheapest price.

Step4. The agent stochastically chooses action a based on the Q-value.

Step5. If $Buy = 1$, the agent purchases M_1 and M_2 from agents who sell them at prices of v_{M_1} and v_{M_2}, respectively.

Step6. All of the producers agents and the consumer agents implement Step2 to Step5.

Step7. All of the producer agents update the Q-values.

Step8. All of the producer agents produce their products using purchased material merchandises. All of the consumer agents consume his merchandise.

Step9. Reaches an end condition when the simulation terminates; otherwise returns to Step2.

We define the flow from Step2 to Step8 as one episode.

V. GINI'S COEFFICIENT

Gini's coefficient (GC) is one of the barometers of social divide in economics, and has a real value between 0 and 1 [15]. When the social divide becomes larger, GC is close to 1. Assuming the presence of n 'non-negative' samples, $y_1, y_2, y_3, ..., y_n (y_1 \leq y_2 \leq y_3 \leq ... \leq y_n)$, then GC is defined as

$$GC = \frac{1}{2\mu_y n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j|, \tag{3}$$

where μ_y represents the average of $y_i (i = 1, 2, 3, ..., n)$. When all of the samples are the same, $GC = 0$. GC is also defined as "The ratio of the area that is surrounded by the Lorenz curve and the line of equal share to the area of the triangle that is under the line of equal share" (Fig.4). The Lorenz curve is described by the summation of the samples to i from 1, that is,

$$\phi_i = \sum_{j=1}^{i} y_j. \tag{4}$$

If we denote that the area that is bounded by the Lorenz curve and the line of equal share as A, the area under the Lorenz curve as B and the area of the triangle that is under
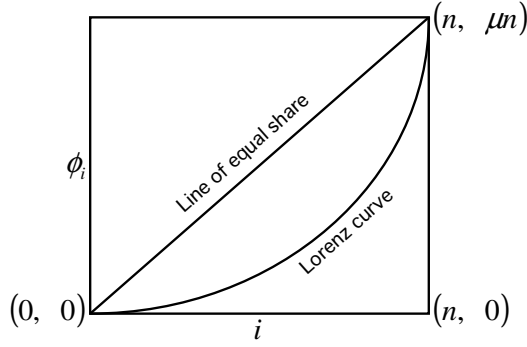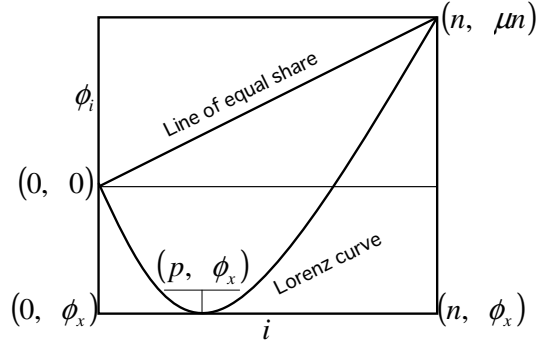
Fig. 4. Lorenz curve.

Fig. 5. Generarized Lorenz curve.

Fig. 6. Relationship between economic efficiency and income divide when the number of consumer agents is 16.

line of equal share as C, then GC is approximated as

$$
\begin{aligned}
GC &= A/C \\
&= (C - B)/C \\
&= 1 - B/C \\
&= 1 - \sum_{i=1}^{n} \phi_i / (\mu_y n^2 / 2) \\
&= 1 - \frac{2}{\mu_y n^2} \sum_{i=1}^{n} \phi_i.
\end{aligned}
\tag{5}
$$

However, GC cannot be used to evaluate the divide correctly when the samples include negative values. In this case, GC has the possibility of a value smaller then 0 or greater than 1. In fact, in the producer economic model considered in this study, the agent's income may be negative. Therefore, we need to extend GC to evaluate the divide emerging in our model. Assuming n samples, $y_1, y_2, y_3, ..., y_n (y_1 \le y_2 \le y_3 \le ... \le y_n$; y_i corresponds to R in section IV), and the samples from y_i to y_x take negative values. Then, the minimum of the Lorenz curve is,

$$
\phi_x = \sum_{i=1}^{x} y_i.
\tag{6}
$$

We redefine GC in an extended description from a geometric view-point, i.e., "The ratio of the area that is surrounded by the Lorenz curve and the line of equal share to the area of a quadrangle defined by $(0,0)$, $(0, \phi_x)$, (n, ϕ_x), $(n, \mu n)$" (Fig.5). Then, the extended Gini's coefficient (GE) can be described as

$$
GE = 1 - \frac{2}{\mu_y n^2 - 2\phi_x n} \sum_{i=1}^{n} (\phi_i - 2\phi_x),
\tag{7}
$$

where GE has a value between 0 and 1 and GE is approximated as

$$
GE = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |z_i - z_j|}{2\mu_z n^2},
\tag{8}
$$

where z_i is the linear transformation of y_i following as
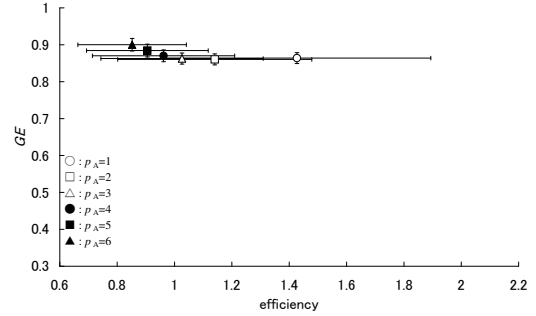
$$
z_i = y_i - \frac{2\phi_x}{n}.
\tag{9}
$$

μ_z represents the average of $z_i (i = 1, 2, 3, ..., n)$. Using y_i, GE is expressed as

$$
GE = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j|}{2(\mu_y - \frac{2\phi_x}{n})n^2}.
\tag{10}
$$

VI. SIMULATION

We simulated 30,000 episodes with the price of mono-bona p_A set from 1 to 6. The number of di-bona producer agents, tri-bona producer agents and tetra-bona producer agents were 64, 32 and 48, respectively. We recorded the average value of GE and the average of economic efficiency from episode 20,000 to 30,000. Because the producer agent achieved sufficient learning, the process was terminated at episode 20,000. Economic efficiency is defined as $\sum_N I / \sum_N C$, where N is the number of producer agents, the learning-rate is $\beta = 0.1$, and the temperature parameter T is 1.0.

Fig. 6 , Fig. 7 and Fig. 8 corresponding to the cases in which the number of consumer agents were 16, 32, 48, respectively, show the average of GE and the economic efficiency. The abscissas and ordinates represent the economic efficiency and GE, respectively. These figures show that economic efficiency decreases with increasing price of mono-bona p_A. This is because the maximum price that can be asked by the producer agent is fixed, in spite of increasing p_A. The producer agents,
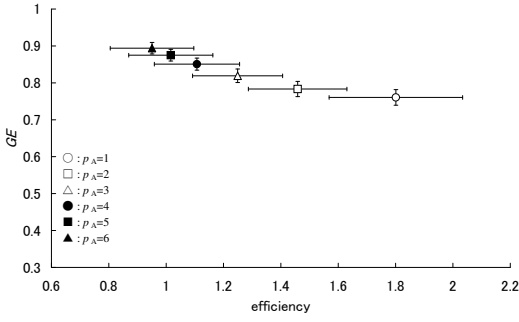
Fig. 7. Relationship between economic efficiency and income divide when the number of consumer agents is 32.
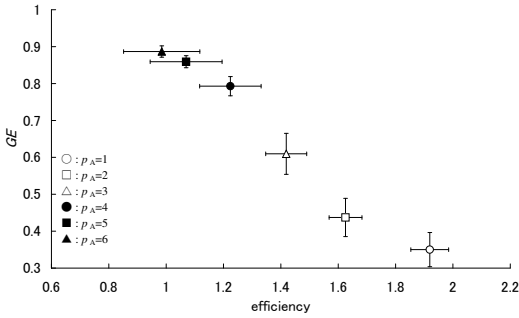

Fig. 8. Relationship between economic efficiency and income divides when the number of consumer agents is 48.

VII. CONCLUSION

In this study, we proposed a producer economic model to evaluate the relationship between economic efficiency and income divide. The results is suggested that for economics under conditions of low competitiveness, being a state composed of relatively more consumer agents than producer agents, display a higher negative relationship between Gini's coefficient and the economic efficiency than those operating under high competitiveness. In addition, the dynamics of the income divide yields an S-shaped curve for the change in economic efficiency when the supply corresponds to the demand.

ACKNOWLEDGMENT

REFERENCES

[1] Terano,T., Kita,H., Kaneda,T., Arai,K., Deguchi,H.: Agent-Based Simulation -From Modeling Methodologies to Real-World Applications- ; Springer-Verlag Tokyo (2005).
[2] Adelinde M. Uhrmacher and Danny Weyns: Multi-Agent Systems - Simulation and Applications- ; CRC Press Taylor & Francis Group (2009).
[3] K,Takadama: Multiagent Learning -Exploring Potentials Embedded in Interation among Agents- ; CORONA PUBLISHING CO.,LTD (2003).
[4] Watkins, C.J.C.H. : Learning from Delayed Rewards; PhD thesis, Cambridge University, Cambridge, England (1989).
[5] Sutton, R. S., Brato, A. G. : Reinforcement Learning; MIT Press, Cambridge, MA, USA, 1998.
[6] Sutton, R. S. : Learning to predict by the methods of temporal difference; Machine Learning,Vol. 3, pp.9-44, 1988.
[7] Kimura, H., Kobayashi, S. : An analysis of actor/critic algorithms using eligibility traces: Reinforcement learning with imperfect value function; In International Conference on Machine Learning, pp.278-286, 1998.
[8] Konda, V. R., Tsitsiklis, J. N. : Actor-critic algorithms; SIAM Journal on Control and Optimization, Vol.42, pp.1143-1146, 2003.
[9] Kkade, S. : A natural policy gradient; In Advances in Neural Information Processing Systems, 14, pp.1531-1538, MIT Press, 2002.
[10] Peters, J., Vijayakumar, S., Schaal, S. : Reinforcement learning for humanoid robotics; In Third IEEE/RAS International Conference on Humanoid Robots, 2003.
[11] Watkins, C. J. C. H., Dayan, P.: Technical Note: Q-Learning; Machine Learning 8, pp.279-292 (1992).
[12] Arai,S.,Miyazaki,K.,Kobayashi,S.: Methodology in Multi-Agent Reinforcement Learning - Approaches by Q-Learning and Profit Sharing-; Technical Papers, Journal of Japanese Society for Artificial Intelligence, Vol.13, No.4, pp.609-618 (1998).
[13] Arai,S.,Miyazaki,K.,Kobayashi,S.: Cranes Control Using Multi-agent Reinforcement Learning; International Conference on Intelligent Autonomous System 5, pp.335-342 (1998).
[14] Sutton, R. S., Barto, A.: Reinforcement Learning; An Introduction, A Bradford Book, The MIT Press (1998).
[15] Kimura,K.: A micro-macro linkage in the measurement of inequality: another look at the Gini coefficient; Quality and Quantity 28:pp.83-97 (1994).

who produce the product made from mono-bona, can increase the price of the product. The tri-bona producer agents increase the price more. The tetra-bona producer agents can-not price above 30, which is the upper limit of the price, so the tetra-bona producer agents' incomes decrease. Therefore, economic efficiency decreases. Economic efficiency is increased with increasing the number of consumer agents. This is because of the increased demand with increasing number of consumer agents, and the producer agents can sell their product at higher prices.

These figures show that GE increased with increasing p_A. In particular, when the numbers of consumer agents is 48, GE exhibits a greater change than in the remaining cases. In the cases where the number of consumer agents are 16 and 32, numbers of the consumer agents are less than those of the tetra-bona producer agents. Therefore, a price-cutting war forces GE higher. On the other hand, when the number of consumer agents is 48, the number of consumer agents and tetra-bona producer agents are the same, so the price-cutting war does not occur. Therefore, GE depends on p_A.

As a result, it is implied that economic efficiency will be reduced by increasing the price of primordial merchandise, and the income divide is increased with reducing economic efficiency. Most notably, when the number of producers and consumers are the same, i.e., when the supply and the demand are the same, the dynamic of the income divide shows an S-shaped curve for the change in the economic efficiency.