Eur. Phys. J. B **47**, 587–593 (2005)
DOI: 10.1140/epjb/e2005-00350-9

THE EUROPEAN PHYSICAL JOURNAL B

Global optimization of minority game by intelligent agents

Yan-Bo Xie[1], Bing-Hong Wang[1,a], Chin-Kun Hu[2], and Tao Zhou[1]

[1] Nonlinear Science Center and Department of Modern Physics, University of Science and Technology of China, Hefei, 230026, P.R. China
[2] Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan

Received 18 June 2005 / Received in final form 20 June 2005
Published online 17 November 2005 –

Abstract. We propose a new model of minority game with intelligent agents who use trail and error method to make a choice such that the standard deviation σ^2 and the total loss in this model reach the theoretical minimum values in the long time limit and the global optimization of the system is reached. This suggests that the economic systems can self-organize into a highly optimized state by agents who make decisions based on inductive thinking, limited knowledge, and capabilities. When other kinds of agents are also present, the simulation results and analytic calculations show that the intelligent agent can gain profits from producers and are much more competent than the noise traders and conventional agents in original minority games proposed by Challet and Zhang.

PACS. 02.50.Le Decision theory and game theory – 05.65.+b Self-organized systems – 87.23.Ge Dynamics of social systems – 87.23.Kg Dynamics of evolution

1 Introduction

The minority game (MG) models were introduced by Challet and Zhang in 1997–1998 for modelling the competition for limited resources [1], which have attracted much attention in recent years. The basic scenario is easy to explain: there is a population of N players who, at each time step, have to choose either 0 or 1. Those who are in the minority side win, the other lose (to avoid ambiguities, N is chosen to be odd). The agents make their decisions based on the most recent m outcomes, thus there are 2^m different histories. A strategy is defined as a table of 2^m choices (either 0 or 1) for the 2^m corresponding histories, so that there are 2^{2^m} different strategies in the strategy-space. Each agent randomly picks $s > 1$ strategies from the strategy-space in the beginning of the MG. To each strategy is associated an integral point, which initially takes the value 0 and will increase by 1 at each time step if it predicts the result correctly. Each agent uses the one with the highest point among his s strategies; if there are several strategies with the same highest point, one of those will be chosen randomly. A very important quantity in this model is the overall loss defined as

$$L(t) = N_{\rm loss}(t) - N_{\rm win}(t) \geq 1 \tag{1}$$

where $N_{\rm loss}$ and $N_{\rm win}$ are, respectively, the number of losers and winners at time t. The smaller $L(t)$ is, the less the overall loss is and thus the better the system performs. Notice that the minimum value of $L(t)$ is 1 when $N_{\rm loss} = (N+1)/2$ and $N_{\rm win} = (N-1)/2$. Another related quantity is called the standard deviation and defined as

$$\sigma^2(t) = (n_0(t) - \bar{n})^2 \tag{2}$$

where n_0 is the number of agents who choose 0 and $\bar{n} = N/2$. It is easy to see that $\sigma^2(t) = L^2(t)/4$ and theoretically, the minimum value of $\sigma^2(t)$ is 0.25.

One of the focuses of scientists' attention is the problem about how to improve the performance of system, i.e. to reduce σ^2. Recently, some new kinds of agent are introduced [2,3], by whom the overall performance of system is improved. A further question is whether it is possible to achieve the global optimization in the framework of the MG model assuming that agents try to outsmart each other for their selfish gain and act based on inductive thinking [4].

Recently, a significant work was achieved by Reents et al., who proposed a stochastic minority game model in which σ^2 is minimized [5]. In their model, an agent will not change his choice in the next time step if he wins in the present turn; on contrary, he will change his choice with a probability p. The value of p is the same for all the agents. When $p \sim 1/N$, Reents et al., found that $\sigma^2 \sim 1$. However, the agents in real-life systems are not as clever as Reents et al. proposed, The agents do not know how to select a value of p, and even do not know the total number of agents N. Thus Reents's model may be not proper for the systems consisting of agent with inductive thinking.

[a] e-mail: bhwang@ustc.edu.cn

Metzler and Horn introduced the evolution into a stochastic MG model [6]. Similar to other evolutionary MG model [7], for an arbitrary agent **i**, a probability $p_i(t)$ and a score s_i is equipped [8]. The score s_i increases by 1 if the agent wins and decreases by 1 if the agent loses. When $s_i \le d < 0$, the agent is deceased and replaced by a new agent with a reset score $s_i = 0$. If $p_i(t)$ of the new agent is randomly distributed in $(0, 1)$, the average value of $p_i(t)$ in the final stationary state has been found to be of the order of 1 and thus $\sigma^2 \sim O(N^2)$. They also discussed the situation in which the new agent chooses $p_i(t)$ by copying the value of $p_j(t)$ of another agent who is randomly selected. Within this scheme, it is possible to see that $p_i(t) \sim O(1/N)$ and $\sigma^2 \sim 1$ in sufficiently long time. However, p is of the order of $1/N$ and σ^2 is greater than 0.25 in the final state. The best solution is still not achieved in their model.

In 2001 Marsili et al. [9] proposed a MG model to reduce σ^2. In their model [9], the agents are adaptive in the sense that they learn from past experience. Explicitly, the ith agent uses the information accumulated in $\Delta_i(t)$ to take decisions ($a_i(t) = 1$ or -1) with the probability

$$\mathrm{Prob}(a_i(t) = \pm 1) = \frac{e^{\pm \Delta_i(t)}}{e^{\Delta_i(t)} + e^{-\Delta_i(t)}}$$

and $\Delta_i(t)$ is updated by

$$\Delta_i(t+1) = \Delta_i(t) - \frac{\Gamma}{N}[A(t) - \eta a_i(t)]$$

where $A(t) = \sum_i a_i(t)$, and Γ and η are the reactive constant and the impact constant, respectively. When the agents take into account their impact on the game $\eta = 1$, minimization of σ^2 is also found for certain range of parameter Γ. However, it can be shown that if the agents are very reactive to the past experience, the oscillatory solution which gives $\sigma^2 \sim N^2$ is firstly reached and stays almost unchanged for very long time even if the agents account for their impact [10]. Since there is no guidance for the agents to choose the reactive constant, i.e., the learning rate in reference [9], the action of those agents in their model should not be classified into the category of inductive thinking.

In the present paper, we propose a new MG model with intelligent agents such that the standard deviation σ^2 and the total loss in this model reach the theoretical minimum values in the limit of long time. The intelligent agents act based on inductive thinking, but are able to bring global optimization to the system. Simulation results and analytic calculations show that when other kinds of agents are also present, the intelligent agent can gain profits from producers (to be defined below) and are much more competent than the noise traders and conventional agents in the original MG models [1].

We should mention that when σ^2 is minimized, a state of Nash equilibrium can be reached [11]. However, Nash equilibrium is a more general concept because it does not imply the minimization of σ^2 [9].

2 Model and numerical simulation

Our model consists of N agents with N an odd integer. Each agent has only one strategy which evolves with the following rule: suppose at a given time step t, the memory (history) is μ and the strategy of the ith agent is $s_i(t, \nu)$ for $\nu = 0, ..., 2^m - 1$. Also, the ith agent has a probability function $p_i(t, \nu)$ for $\nu = 0, ..., 2^m - 1$. If the ith agent wins at t, the strategy will not be changed; contrarily, with probability $1 - p_i(t, \mu)$, all of $s_i(t, \nu)$ are not changed, with probability $p_i(t, \mu)$, $s_i(t+1, \mu) = 1 - s_i(t, \mu)$, but $s_i(t+1, \nu) = s_i(t, \nu)$ for all other $\nu \ne \mu$.

The initial value of $p_i(t, \nu)$ is randomly selected in $(0, 1)$ and evolves by self-teaching mechanism, which is the simplest trail and error method. For a given time step t with history μ, consider the last time step t' when the memory is also μ. If the agent **i** won at t' or he lost but did not change $s_i(t', \mu)$, then no changes aiming at $p_i(t, \mu)$ will occur. Otherwise, $p_i(t, \mu)$ will change according to the following rule:

$$p_i(t+1, \mu) = \begin{cases} \min(1, 2p_i(t, \mu)) & \text{agent } i \text{ wins at time } t \\ p_i(t, \mu)/2 & \text{agent } i \text{ loses at time } t. \end{cases} \tag{3}$$

No changes will occur for all $p_i(t, \nu)$ with $\nu \ne \mu$.

Let us explain our model in a few more words. For simplicity, consider the case $m = 0$ and thus the history index ν can be dropped. In this case, the model can be described by $s_i(t)$ and $p_i(t)$ at time t. Suppose the ith agent wins at the time step $t-1$, then $s_i(t) = s_i(t-1)$. Suppose the ith agent loses at $t-1$, then with probability $1 - p_i(t-1)$, $s_i(t) = s_i(t-1)$; with probability $p_i(t-1)$, $s_i(t) = 1 - s_i(t-1)$. The evolution of $p_i(t)$ is given as follows. If $s_i(t) = s_i(t-1)$, then $p_i(t+1) = p_i(t)$. Only when $s_i(t) = 1 - s_i(t-1)$, i.e., when the ith agent loses at $t-1$ and changes his strategy s_i at t, $p_i(t+1)$ may be different from $p_i(t)$ according to equation (3) based on his performance at t.

The formula (3) can be understood easily. If agent **i** lost at $t-1$ and changed $s_i(t) = 1 - s_i(t-1)$ at time t, and he wins at time t, he will think that to change strategy when he loses may be advisable and therefore $p_i(t+1)$ is increased (This implies that he will change his action more frequent when he loses at later time.); otherwise, if he lost at $t-1$ and changed $s_i(t) = 1 - s_i(t-1)$ at time t, and he loses again at time t, then he may think that the change was too hurried, therefore $p_i(t+1, \mu)$ is reduced.

We present the simulation results in Figure 1, which shows that the system will reach global optimization in sufficiently long time. We have checked the behaviors of time evolution of $\sigma^2(t)$ for the cases with more agents and larger memory sizes; the results are the same as that of $N = 101$ and $m = 0,\ 1,\ 2$.

Figure 2 presents the log-log plot for the time dependence of $G(t) = \sum_{i=1}^{N} p_i(t)$ for $N = 101$ and $m = 0$. The results show that $G(t)$ has a power law dependence on time with the exponent $\gamma \approx -1$ when t is large, which suggests that $G(t) \to 0 (t \to \infty)$, thus it is reasonable to suppose $p_i(t) \ll 1/N$ when t is sufficiently large. In this case, at most one agent may change the strategy at each

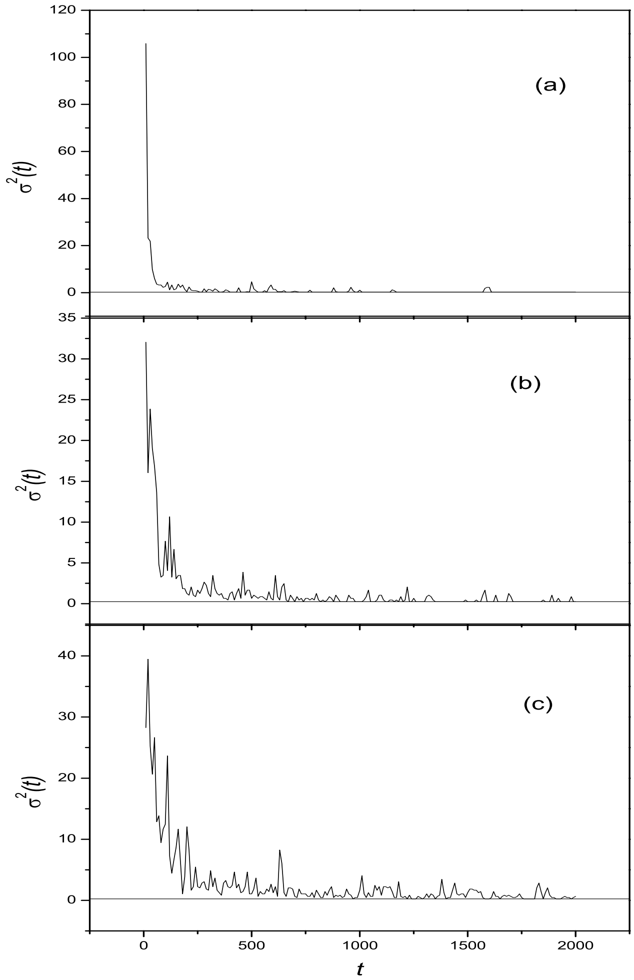

Fig. 1. Time evolution of $\sigma^2(t)$ for $N = 101$ intelligent agents with m =0 (a), 1 (b), and 2 (c). The value of $\sigma^2(t)$ shown in this figure is the average of 10 independent simulations and the horizontal line represents $\sigma^2 = 0.25$.

time step (the probability for two or more agents changing their strategies at the same time is negligibly small) thus the number of agents on the majority side is always $(N + 1)/2$. Therefore, the agent who changes the strategy is from the losing side to the losing side and $p_i(t)$ is reduced by a factor of 2. Since $p_i(t) \ll 1/N$, the probability that one agent will change his strategy is

$$\eta = 1 - \prod_{i \in W_l(t)} (1 - p_i(t)) \approx \sum_{i \in W_l(t)} p_i(t) \approx \frac{G(t)}{2}$$

where $W_l(t)$ is the set of losers at time t. Then we have the iterative equations for $G(t)$:

$$G(t + 1) = \eta \frac{2N - 1}{2N} G(t) + (1 - \eta)G(t). \quad (4)$$

According to equation (4), one can find that $G(t) \sim t^{-1}$ [12], which is consistent with the simulation result shown in Figure 2.

It may be helpful to mention that roughly N time steps are required for our system to reach the theoretical minimum $\sigma^2 \sim 0.25$. Therefore, when N is very large, very long time are required to reach the global optimization.

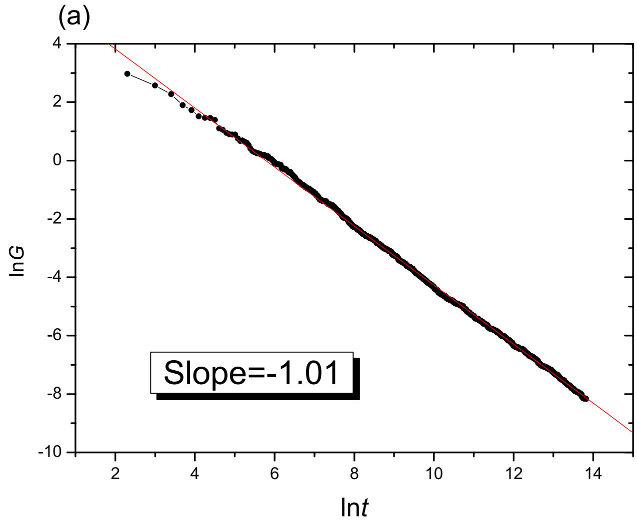

Fig. 2. Time dependence of $G(t)$, where $N = 101$ and $m = 0$. The slope of the curve in this figure is $-1.01(\approx -1)$.

3 Intelligent agents in mixed market

Challet et al. classified the agents into three different types [13]: producers who have only one fixed strategy for a given history, speculators (conventional agents in original minority game) who have two or more strategies, and the noise traders who make their choices by random tosses. In this section, we will investigate how intelligent agents perform in mixed market [14].

Note that the evolution of $p_i(t, \mu)$ for different memories is essentially decoupled in our model when only intelligent agents are present. Therefore, mathematically speaking, the $m \neq 0$ case is a trivial generalization of the $m = 0$ case in this situation. The situations change greatly when other types of agents are present. In these situations, the $m \neq 0$ case will not be the trivial generalization of the $m = 0$ case.

Firstly, let us look into how the intelligent agents compete with the producers. Assume that there are N_p producers and N_s intelligent agents with $N_p + N_s$ an odd integer, each producer has only one fixed strategy. For simplicity, we shall first discuss the case of $m = 0$. Suppose N_{p0} producers always choose 0, and N_{p1} producers always choose 1. If $\Delta = N_{p0} - N_{p1} > N_s(< -N_s)$, then N_s intelligent agents must choose 1(0) in the equilibrium state and win at each time step. When $N_s > \Delta > 0$ (the case $N_s > -\Delta > 0$ is analogic), the situation is slightly complicated. From the discussion in Section 2, it is not difficult to see that the overall loss of $N_p + N_s$ agents is minimized in the equilibrium state. Namely, there will be either $(N_s - \Delta + 1)/2$ intelligent agents choosing 0 and $(N_s + \Delta - 1)/2$ intelligent agents choosing 1 or $(N_s - \Delta - 1)/2$ intelligent agents choosing 0 and $(N_s + \Delta + 1)/2$ intelligent agents choosing 1. In the former case, the agents choosing 0 are losers, while in the latter case, the agents choosing 0 are winners. The equilibrium state is described by the transition between two cases. Before it switches to another case, the equilibrium state stays in one case for a period of time, called the life time. The life times of two cases are different. Assume that the distribution of probability p_i of agent **i** has a sharp peak around $\langle p\rangle$, then the life time of the former case is $\tau_1 = 2/(N_s - \Delta + 1)\langle p\rangle$ and the latter case is $\tau_2 = 2/(N_s + \Delta + 1)\langle p\rangle$, where $\langle p\rangle$ denotes the average value of p_i. The overall gain of the intelligent agents at each time step is equal to

$$\begin{aligned}\Sigma =&\frac{1}{\tau_1+\tau_2}[(\frac{N_s+\Delta-1}{2}-\frac{N_s-\Delta+1}{2})\tau_1\\&+(\frac{N_s-\Delta-1}{2}-\frac{N_s+\Delta+1}{2})\tau_2]\\=&\frac{1}{N_s+1}[\Delta^2-1-N_s]. \qquad (5)\end{aligned}$$

Therefore, $\Sigma > 0$ when $\Delta < N_s < \Delta^2 - 1$. The average profit gained by each intelligent agent at each time step is

$$\frac{\Sigma}{N_s} = \frac{1}{N_s(N_s+1)}[\Delta^2 - 1 - N_s]. \qquad (6)$$

According to equation (6), when $N_s < \Delta^2 - 1$, each intelligent agent can gain profits from producers. Suppose the number of intelligent agent N_s is not fixed, if $N_s < \Delta^2 - 1$, some new intelligent agents, if available, will join the game since they can gain profits from producers. Thus there will be eventually $N_s \approx \Delta^2 - 1$ intelligent agents in the market, whose profits are approximatively equal to 0 with small fluctuations. This process can be considered as an example for the efficient market hypothesis (EMH), which is hotly controversial in the recent years [15]. But in real-life financial market, the number of producers is not fixed, thus the equilibrium state can rarely be reached.

When $m > 0$, the number of possible histories is $2^m > 1$. For a given history μ, suppose $N_{p0}(\mu)$ producers always choose 0 and $N_{p1}(\mu)$ producers always choose 1. Then $\Delta(\mu) = N_{p0}(\mu) - N_{p1}(\mu)$ is a function of μ. Since different history μ is essentially decoupled in our model and the number of intelligent agents N_s is fixed, there may be three cases under history μ: (i) $|\Delta(\mu)| \geq N_s$, each intelligent agent can gain one point at each time step; (ii) $\Delta^2(\mu) - 1 > N_s > |\Delta(\mu)|$, the intelligent agents can averagely gain profit from the producers; (iii) $\Delta^2(\mu) - 1 < N_s$, the intelligent agent cannot gain profit and are characterized by the overall loss described by equation (1).

The above picture is confirmed by the numerical simulation result shown in Figure 3a. One can find that σ^2 decreases as t increases and decays to 0.25 when t is sufficiently large. Figure 3b plots the time dependence of the mean gain for intelligent agents:

$$A_s(t) = \frac{N_{s\mathrm{win}}(t) - N_{s\mathrm{lose}}(t)}{N_s}$$

where $N_{s\mathrm{win}}$ and $N_{s\mathrm{lose}}$ denote the number of intelligent agents who win and lose, respectively. Initially, $A_s(t)$ is negative, but as t increases, $A_s(t)$ becomes positive. Therefore, the intelligent agents can gain profits from producers in the regime $\Delta^2(\mu) - 1 > N_s$.

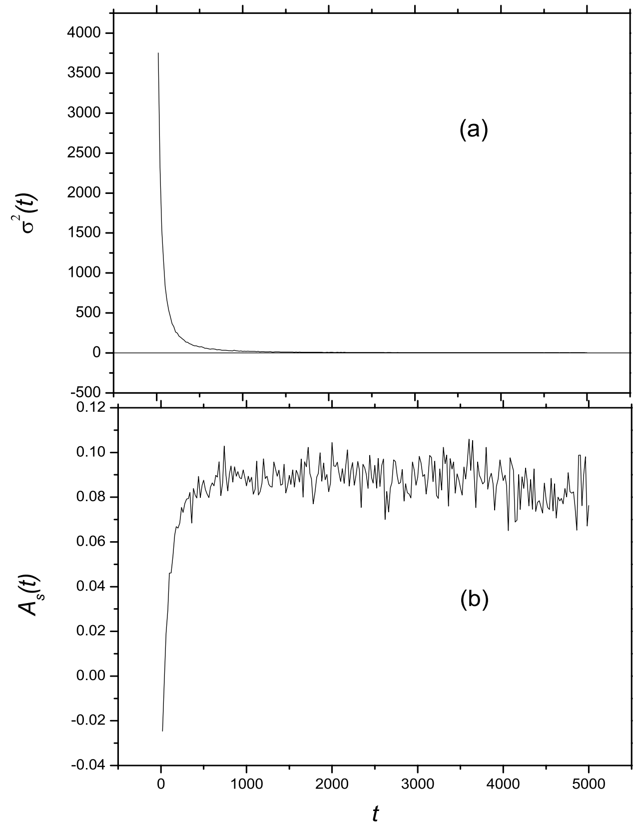

Fig. 3. Time evolution of $\sigma^2(t)$ (a) and $A_s(t)$ (b), where $N_p = 200$, $N_s = 801$, $m = 1$ and $\Delta(0) = \Delta(1) = 200$. The value of $\sigma^2(t)$ and $A_s(t)$ shown in these two figures is the average of 32 independent simulations and the horizontal line in figure (a) represents $\sigma^2 = 0.25$.

Secondly, let's consider the case in which the noise traders and intelligent agents are present. Assume that there are N_n noise traders and N_s intelligent agents with $N_n + N_s$ an odd integer. Figure 4a plots the time dependence of σ^2, one can find that σ^2 decreases as t increases, but does not reach the theoretical optimal 0.25 in the limit of long time. This result is not difficult to understand for the existence of noise traders will bring more fluctuations into the system. Figure 4b and 4c exhibit the time dependence of A_s and A_n respectively, where A_n is the mean gain of noise traders:

$$A_n(t) = \frac{N_{n\mathrm{win}}(t) - N_{n\mathrm{lose}}(t)}{N_n}$$

$N_{n\mathrm{win}}$ and $N_{n\mathrm{lose}}$ denote the number of the noise traders who win and lose, respectively. Apparently, the intelligent agents perform much better than the noise traders do.

We have also studied the case in which the conventional agents, who take the actions based on the original minority game model [1] with the memory size m_m, and intelligent agents, who take the $m = m_s = 0$ case, are present. Assume that there are N_s intelligent agents and

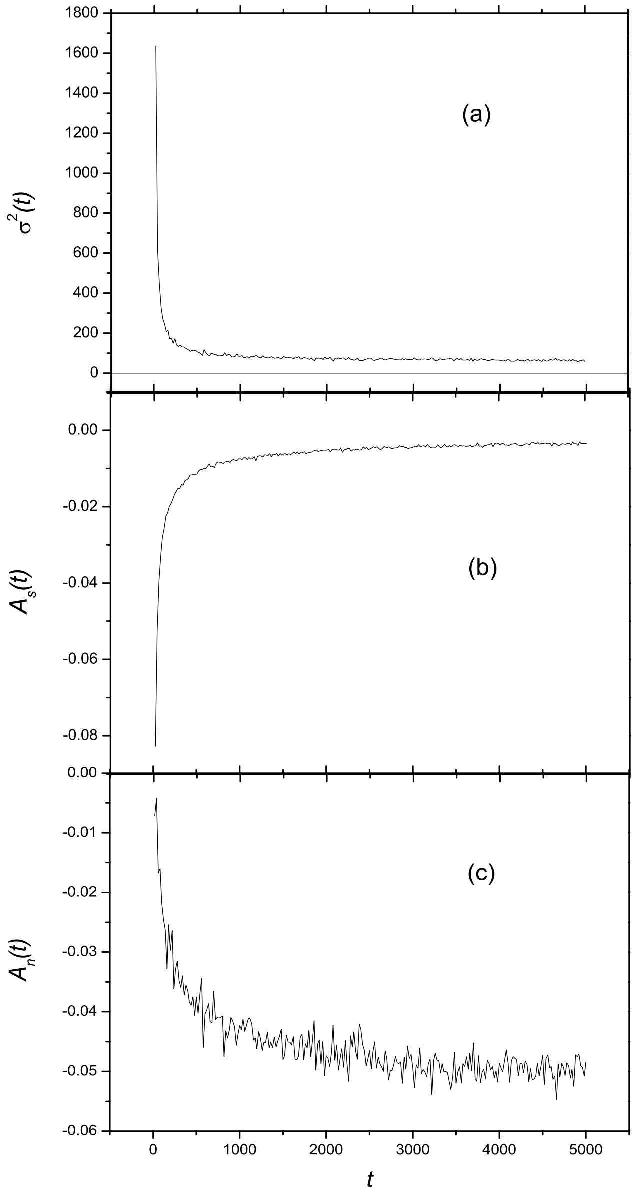

Fig. 4. Time evolution of $\sigma^2(t)$ (a), $A_s(t)$ (b), and $A_n(t)$ (c), where $N_s = 801$, $N_n = 200$ and $m = 1$. The value of $\sigma^2(t)$, $A_s(t)$ and $A_n(t)$ shown in these three figures is the average of 32 independent simulations and the horizontal line in figure (a) represents $\sigma^2 = 0.25$.

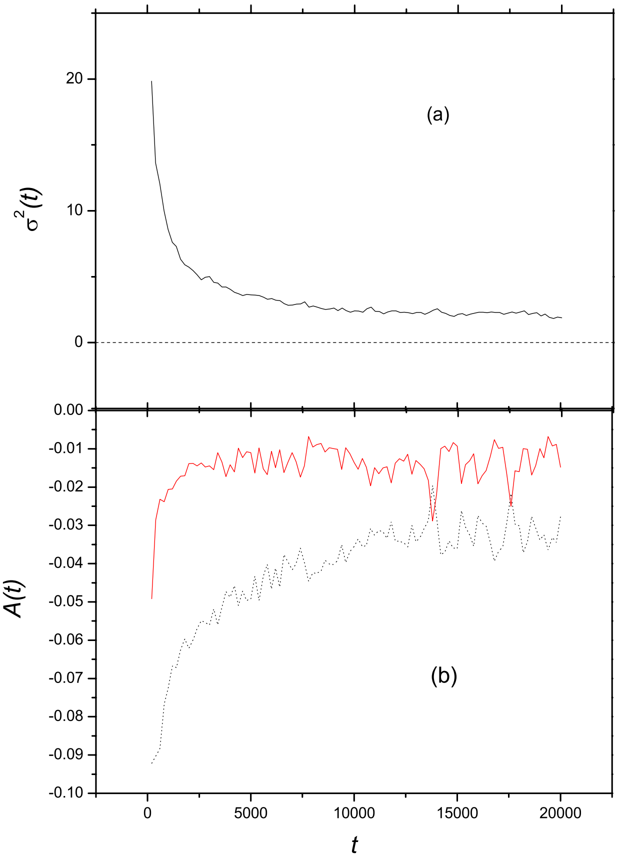

Fig. 5. Time evolution of $\sigma^2(t)$ (a), $A_s(t)$ and $A_m(t)$ (b). In Figure 5b, the dotted line represents the results of $A_m(t)$, the solid line represents the results of $A_s(t)$. The parameters for the intelligent agents are $N_s = 51$ and $m_s = 0$. The parameters for the conventional agents are $N_m = 50$, $m_m = 3$ and the number of strategies $S = 2$. The value of $\sigma^2(t)$, $A_s(t)$ and $A_m(t)$ shown in these two figures is the average of 32 independent simulations and the horizontal line in figure (a) represents $\sigma^2 = 0.25$.

N_m conventional agents with N_s+N_m an odd integer. Figure 5 presents our numerical studies for the case $m_m = 3$. Figure 5a shows the time dependence of σ^2. One sees that σ^2 decreases with time but also does not reach the theoretical optimal value 0.25 in the limit of long time. This result implies that the conventional agents also introduce fluctuations into the system, though its magnitude is less than those of the noise traders. In Figure 5b, we report the time dependence of A_s and A_m respectively, where A_m is the mean gain of conventional agents:

$$A_m(t) = \frac{N_{m\text{win}}(t) - N_{m\text{lose}}(t)}{N_m}$$

where $N_{m\text{win}}$ and $N_{m\text{lose}}$ are the number of the conventional agents who win and lose, respectively. From Figure 5b, one immediately finds that the intelligent agents perform much better than the conventional agents. Figure 6 presents our numerical studies for the case $m_m = 0$. Comparing with Figure 5, one finds that the results for these two cases are almost the same.

It may be interesting to ask the question how the presence of the intelligent agents changes the performance of conventional agents in the original MG model [1]. We have also done the numerical simulations for the parameters $N_s = 0$, $N_m = 51$ and $m_m = 3$. We have found that $A_m(t)$ approaches -0.07, which is below the value -0.035 when the intelligent agents are present and $t \sim 20000$ (Fig. 6). This implies that the performance of the conventional agents is improved when the intelligent agents are introduced into the model.

Finally, we have studied the case in which the adaptive agents in the work by Marsili et al. [9] mix with the intelligent agents introduced in this paper. Figure 7 presents

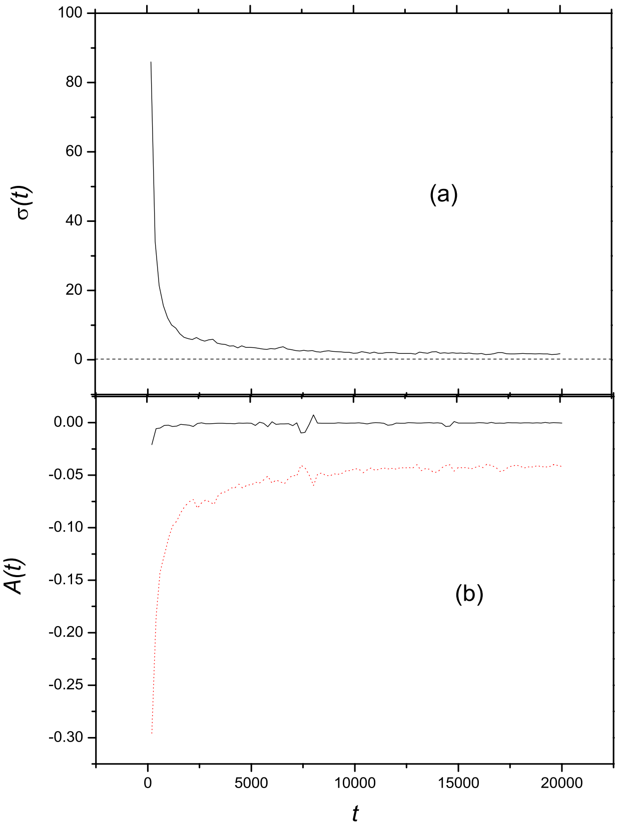

Fig. 6. Time evolution of $\sigma^2(t)$ (a), $A_s(t)$ and $A_m(t)$ (b). In Figure 6b, the dotted line represents the results of $A_m(t)$, the solid line represents the results of $A_s(t)$. The parameters for the intelligent agents are $N_s = 51$ and $m_s = 0$. The parameters for the conventional agents are $N_m = 50$, $m_m = 0$ and the number of strategies $S = 2$. The value of $\sigma^2(t)$, $A_s(t)$ and $A_m(t)$ shown in these two figures is the average of 32 independent simulations and the horizontal line in figure (a) represents $\sigma^2 = 0.25$.

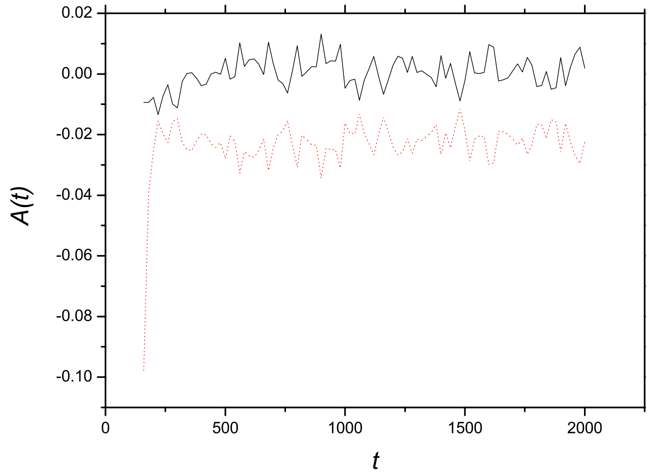

Fig. 7. Time evolution of the mean gain for the adaptive agents $A_m(t)$ and for the intelligent agents $A_s(t)$. The dotted line represents the results of $A_m(t)$, the solid line represents the results of $A_s(t)$. The parameters for the intelligent agents are $N_s = 51$ and $m = 0$. The parameters for the adaptive agents are $N_m = 50$, $\Gamma = 5$, and the impact parameter $\eta = 1$. The value of $A_s(t)$ and $A_m(t)$ shown in these two figures is the average of 320 independent simulations.

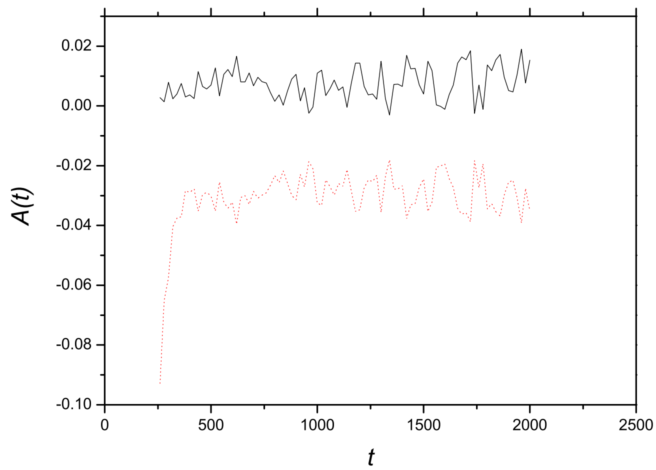

Fig. 8. Time evolution of the mean gain for the adaptive agents $A_m(t)$ and for the intelligent agents $A_s(t)$. The dotted line represents the results of $A_m(t)$, the solid line represents the results of $A_s(t)$. The parameters for the intelligent agents are $N_s = 51$ and $m = 0$. The parameters for the adaptive agents are $N_m = 50$, $\Gamma = 20$, and the impact parameter $\eta = 1$. The value of $A_s(t)$ and $A_m(t)$ shown in these two figures is the average of 320 independent simulations.

the results for the case that the reactive parameter $\Gamma = 5$. One finds that the adaptive agents perform significantly worse than the intelligent agents do. Figure 8 presents the results for the case $\Gamma = 20$ which is essentially identical to the case $\Gamma = 5$. It is interesting to remark that the oscillatory state, which is observed when only the adaptive agents are included for the case $\Gamma = 20$, is now removed by the presence of the intelligent agents.

4 Discussion and conclusion

We propose a new MG model with intelligent agents, who use trail and error method to make a choice. When only the intelligent agents are present, the overall standard deviation (and hence loss $L(t)$ of Eq. (1)) is minimized to the theoretical lower limit $\sigma^2 \to 0.25$ as $t \to \infty$. Notice that although such intelligent agents are independent and only trying to do their best for their selfish gain based on inductive thinking, the Global Optimization is achieved in our model. The result suggests that an economic systems can be self-organized into a highly optimal state by agents who make decisions based on inductive thinking using their limited knowledge and capabilities.

In mixed market cases, when the model consists of the intelligent agents and the producers with only one fixed strategy, we have found that, under certain circumstances, the intelligent agents can gain profit from the producers. Also, the overall loss of the producers and the intelligent agents is minimized. When the model consists of the intelligent agents and the noise traders who choose the room randomly at each round, it is found that the intelligent agents also cooperate very well so that the overall loss of the intelligent agents becomes very small when the time is sufficiently large.

It is worthwhile to emphasize that, the intelligent agents perform much better than the conventional agents in mixed market. Imagine an agent trying to figure out the regularity of the financial market. Assume at time t_1, he has the selection rules for all possible histories, i.e., he has a strategy. At a later time t_2, he finds that the selection rules for some histories do not give profits. Therefore, he may change the selection rule for these histories, but not for the other histories which still give him profits. This is in contrast with the original MG model in which an agents selects the strategy with the highest virtual point. When he changes the strategy, he may change many selection rules although they still make profits. We consider that this is the reason why the conventional agents are less competent than intelligent agents.

This work is supported by the State Key Development Programme of Basic Research of China (973 Project), the National Natural Science Foundation of China under Grant No.70271070, and No.10472116, the Specialized Research Fund for the Doctoral Program of Higher Education of China(SRFDP No.20020358009), the National Science Council (Taiwan) under Grant No. NSC 93-2112-M-001-027, Academia Sinica (Taiwan) under Grant No. AS-91-TP-A02, and National Center for Theoretical Sciences at Taipei.

References

1. D. Challet and Y.-C. Zhang, Physica A **246**, 407 (1997); Y.-C. Zhang, Europhys. News **29**, 51 (1998); D. Challet, Y.-C. Zhang, Physica A **256**, 514 (1998)
2. Y. Li, R. Riolo, S. Savit, Physica A **276**, 234 (2000)
3. H.-J. Quan, B.-H. Wang, P.-M. Hui, Physica A **312**, 619 (2002); H.-J. Quan, B.-H. Wang, P.-M. Hui, X.-S. Luo, Physica A **321**, 300 (2003)
4. H. Simon, *Models of Bounded Rationality* (MIT Press, Cambridge, 1997); W.B. Authur, Amer. Econ. Assoc. Papers Proc. **84**, 406 (1994)
5. G. Reents, R. Metzler, W. Kinzel, Physica A **299**, 253 (2001)
6. R. Metzler, C. Horn, Physica A **329**, 484 (2003)
7. N.F. Johnson, P.-M. Hui, R. Jonson, T.S. Lo, Phys. Rev. Lett. **82**, 3360 (1999); T.S. Lo, S.W. Lim, P.-M. Hu, N.F. Johnson, Physica A **287**, 313 (2000)
8. If agent **i** wins at time t, he will not change his choice, but if he loses, he may change his choice at probability $p_i(t)$
9. M. Marsili, D. Challet, Adv. Complex Systems **4**, 3 (2001); M. Marsili, Physica A **299**, 93 (2001)
10. When the reactive constant is sufficiently large (For instance for the number of agents $N = 101$, the reactive constant $\Gamma = 10$ is sufficiently large.), the oscillatory state is stable for very long time which is much longer than the typical relaxation time in our model (See Figure 1). This can be easily confirmed by numerical simulation
11. R.B. Myerson, *Game Theory: Analysis of Conflict* (Harvard University Press, Cambridge, Massachusetts, 1997)
12. In equation (4), note that $\eta = G(t)/2$, we have $G^{-2}dG = \frac{1}{4N}dt$, then $\int G^{-2}dG \sim \int \frac{1}{4N}dt$, and we can obtain the order of $G(t)$: $G(t) \sim t^{-1}$
13. D. Challet, M. Marsili, Y.-C. Zhang, Physica A **276**, 284 (2000)
14. Mixed market means the financial market consisting of at least two different kinds of agents
15. A. Lo, C. Mackinlay, *A Non-Random Walk Down Wall Street* (Princeton University Press, 1999); S.-H. Chen, C.-H. Yeh, J. Economic Dyamics & Control **25**, 363 (2001); T. Zhou, P.-L. Zhou, B.-H. Wang, Z.-N. Tang, J. Liu, Int. J. Mod. Phys. B **18**, 2697 (2004)