Evolving Specialization, Market and Productivity in an Agent-Based Cooperation Model

Erbo Zhao, Guo Liu, Dan Luo, Xing'ang Xia, and Zhangang Han*

Department of Systems Science, School of Management, Beijing Normal University, Beijing, 100875, People's Republic of China
zhan@bnu.edu.cn

Abstract. This paper introduces an agent-based model in which self-interest intelligent agents are adaptive. Agents can either go to find resources in the environment or mine the resources found. Agents trade information about resources in a market. A biased learning mechanism is introduced to update agents' capabilities of mining and searching. The learning mechanism plays a vital role in the specialization process in our model. Expectation is also introduced in this paper to determine the trade price. Simulations show that agents can specialize in available capabilities, form market and cooperate to increase their wealth. These emergencies come out through just pre-defining some learning and pricing mechanisms that are not so complex but close to reality. Total productivity and market formation are tracked during the evolving process. The wealth distribution during whole evolving process also demonstrates an interesting power law distribution.

Keywords: Agent-based model, individual reinforcement learning, power law, specialization, market forming, expected productivity.

1 Introduction

Agent-based models are increasingly recognized as powerful tools for modeling complex adaptive systems. They are useful in the terms of modeling of the interactions among agents, representation of learning in a dynamic strategy context, and description of the relationships between individual and its local environment [14]. Agent-based models can track individual interactions that otherwise cannot be represented in traditional models that use global variables. In particular, agent-based models excel at relating the heterogeneous behavior of agents with different information, different decision rules, and different situation to the macro behavior of the overall systems [15]. Recently, agent-based models have been applied to a variety of systems, such as biological ecosystems, economics systems [1, 3, 9, 13, 17, 18] and social intelligence systems [4].

* Correspondence author. Tel. 86.10.58807876. This research is supported by National Science Foundation of China under grant No. 60774085 and No. 70601002.

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 1564–1574, 2009.

Agents we are interested in, in this paper, are self-interest individuals that do not have complete information about the whole system. They interact with each other and with the environment to maximize their own profits. On a macro scale, the system composed of these agents may reveal some collective behaviors that are not simple summations of individual properties. A major task of complexity study is to find underlying mechanisms that govern the collective behaviors. The interactions between individual agents and(or) between colonies produce a great many characteristics at the global scale, such as nonlinear oscillation, clustered volatility, and chaos [7, 11, 12], even though the interactions between agents are localized and very simple. Holland argued that the self-adaptive properties of individual agents result in these complexities [8]. Learning is crucial to agent's self-adaptation. Nicolas illustrated that there are essential differences between individual learning and social learning [10]. A biased learning mechanism, which is crucial to the specialization process, is presented to update mining and searching capabilities.

Power law distribution is widely observed in real world systems, such as Pareto's law in economics and Zipf's law in lexicostatistics. For a review of power law distribution, please read ref. [16]. Elsewhere empirical size distributions for which power-law behavior is discovered, including those of sand particle sizes; of meteor impacts on the moon; of numbers of species per genus in flowering plants etc.

We give a model in this paper that simulates a social system as an evolutionary process. This multi-agent system consists of a great number of agents trying to obtain resources in an environment. Agents can take two kinds of tasks: search for resources and exploit the resources found. In the model, an agent can randomly walk to search for resources. When it finds a pack of resources, it may decide to exploit them or sell the information of resources through a market-like mechanism. There is no global planning mechanism in this multi-agent system. Every agent is self-interested, and its goal is to maximize its own profit. Each agent has searching and exploiting capabilities that can be adjusted on the basis of its profit obtained previously.

The model is presented in detail in section 2. The simulated results about specializing process, market forming and wealth distribution are shown in section 3, and conclusion remarks are drawn in section 4.

2 The Model

Consider a system with a great many of resources and M self-interested individual agents. Agents are randomly scattered over a lattice. Each agent obtains profits through exploiting resources or selling the information about discovered resources. During the process of exploiting or selling, the agent ceaselessly studies, improves the skills and makes decisions in the market in accordance with the experiences, to maximize its profit.

2.1 Environment and Resources

The environment is an $L \times L$ lattice with periodic boundary condition. In other words, an agent will step in the lattice from the opposite side when it steps out from one side. An agent only occupies one site, and one locus can only be occupied by no more than one agent at the same time.

Resources are randomly distributed as "packs" over the lattice. Each pack of resource is constant R. The aggregate resources in the system are always maintained as a constant value. When the system is initialized at the beginning of the simulation, and their gross value is always a constant with global density ρ. Once a pack of resources is completely exploited, a new pack will be immediately replenished in any site except for the locus occupied by other packs of resources. As the same as the agents, one pack of resources one locus, and one locus can only be occupied by no more than one pack of resources.

2.2 Neighborhood and Distance

The environment is a lattice, which is a two-dimensional discrete space. In this paper we use Moore neighborhood which is a square-shaped neighborhood surrounding a given site(x_0, y_0). The Moore neighborhood of range r is defined as

$$N^{M}_{(x_0,y_0)} = \{(x, y) : |x - x_0| \le r, |y - y_0| \le r\}. \quad (1)$$

So, we get its area $(2r + 1)^2$ on a square grid.

When referring to the distance between two sites, we have to take into account nine possible distance calculations due to the periodic boundary condition. We definite the minimum distance D_{AB} among the nine distance calculations as the distance between the two points.

$$D_{AB} \triangleq Min\{d_{AB}\}, d_{AB} \triangleq |x_A - x_\beta| + |y_A - y_\beta|, \beta \in \{B, B_1, \ldots, B_8\}. \quad (2)$$

2.3 Agent Properties

Any agent has two kinds of tasks: searching resources and exploiting resources, but it can only participate in one task at each time step. We set up a rule that an agent must exploit its resources which have been discovered, if the resources are not sold at last time step. At each time step, each agent must be in one of the states, mining or foraging. The agents who have got new resources through searching in the environment or buying in the market should be in the state of mining, and those who do not have resources should be in the state of foraging. An agent who does not have resources may be an agent who has sold its information of discovered resources in last time step, or an agent who has to search new resources and did not successfully purchased any resources in the market.

As shown in Fig. 1, an agent can transfer from one state to another. If we express it in a detailed manner, an agent who is exploiting will keep on mining until that pack of resources is exhausted. When an agent has finished exploiting

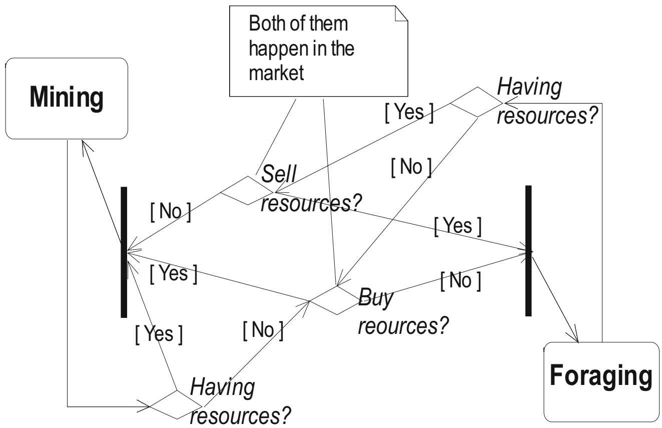

Fig. 1. State-chart of an agent's state transition between the two kinds of states: mining and foraging. An exploiting agent will keep on mining until that pack of resources is exhausted. An agent who has finished exploiting resources or was searching and did not find any resources at the last step goes to the market to purchase resources. An agent goes to the market to sell the information of the resources found if it just found a pack of resources at the last step. A buyer will go searching if it did not successfully purchase any resources. A seller will go searching if it has sold its resources. A buyer will exploit the resources if it successfully purchased a pack of resources. A seller will exploit the resources if it did not successfully sell the resources.

resources, it will go to the market, as a buyer, to buy the information about the discovered resources. If it successfully purchases the information, it will exploit the newly found resources and keep staying in the mining state. If the purchase is not successful, it has to go to the environment to seek new resources, hence its state transfers to the foraging state. One agent always goes to the market, as a seller, to sell the information of the resources just found at the last step. A seller will go searching if it has sold its resources, and keeps in foraging state. Or it must transfer to the state of mining and exploit the resources found by itself if the selling is not successful. One agent who failed to find new resources at last time step goes to the market as a buyer. If it is successful to purchase the information about new resources, it will transfer to the state of mining; or it has to go back to the environment to seek new resources and keep staying in the state of foraging.

Each agent has two kinds of basic capabilities, to search resources and exploit resources. When an agent is searching new resources, it can see the distribution of resources in its neighborhood (we use Moore neighborhood) with the radius of r, including its current site. But it can just randomly choose one pack of unlocked resources, even though it has seen many packs. The searching efficiency is defined K_s as the pack of resources one agent can find in one time step, and we get $K_s = 1 - (1-\rho)^{(2r+1)^2}$, where r is the radius of the Moore neighborhood of an agent's sight and ρ is a constant value of resource distribution density in the environment. The exploiting efficiency is defined K_e as the accumulation of the capability to mine resources in one time step. We can get $K_e = E/E_{max}$. The exploiting capacity E is limited by a maximization number E_{max}. This, we think, is a natural hypothesis that every physical variable should be limited by a finite number.

According to agents' conduct, we divide all agents in our model into three different types: agent in autarky, miner and searcher. An agent is in autarky if it exploits the resources found by itself at time step t and gets profit $R_a(t)$. An agent is a miner if it exploits the resources purchased in the market and gets profit $R_e(t)$, which is the total value of the pack of resources minus the price paid at the market and a tax (see next section). A searcher is an agent which can successfully sell the information of resources it found and gets profit $R_{s(t)}$ which is the price paid by the buyer.

In addition, *movement cost* C is introduced as

$$C = D \times c, \tag{3}$$

where, D is the distance that the agent move from one site to another according to the definition (2), and c is the unit cost that an agent spends in order to cover a unit distance. The costs are calculated proportional to the ratios of wealth accumulation methods.

$$\begin{array}{l} C_e(t) = C \cdot W_e(t-1)/W(t-1), t > 0 \\ C_s(t) = C \cdot W_s(t-1)/W(t-1), t > 0 \\ C_a(t) = C \cdot W_a(t-1)/W(t-1), t > 0 \end{array} \tag{4}$$

$W_e(t)$ is the agent's wealth got from being a miner until time step t, and $W_s(t)$ as a searcher, $W_a(t)$ as in autarky respectively. $W(t)$ is the sum of all kinds of wealth. The initial wealth R is evenly distributed to the three kinds of wealth. So, we get an agent's wealth at time step t

$$\begin{array}{r} W(t) = W_e(t) + W_s(t) + W_a(t) \\ W_e(0) = W_s(0) = W_a(0) = R/3 \\ W_e(t) = W_e(t-1) + R_e(t) - C_e(t), t > 0 \\ W_s(t) = W_s(t-1) + R_s(t) - C_s(t), t > 0 \\ W_a(t) = W_a(t-1) + R_a(t) - C_a(t), t > 0 \end{array} \tag{5}$$

The increasing revenue ΔW, which an agent obtains in the time interval τ, may be divided into two parts: ΔW_s and ΔW_e. The former is the searching revenue at time interval τ resulted from payments of selling the resource information, and the latter is the exploiting revenue directly got by exploiting resources purchased. But the obtained revenue will be divided half into ΔW_s and half into ΔW_e if an agent is in autarky. So, we get

$$\begin{array}{l} \Delta W = \Delta W_e + \Delta W_s \\ \Delta W_e = \Delta W_e(t) - \Delta W_e(t-\tau) + \frac{1}{2}\left(W_a(t) - W_a(t-\tau)\right) \\ \Delta W_s = \Delta W_s(t) - \Delta W_s(t-\tau) + \frac{1}{2}\left(W_a(t) - W_a(t-\tau)\right) \end{array} \tag{6}$$

Agents learn from their history. An agent can gain some useful experience from its previous profits and tasks, modify its corresponding capabilities to pursuit more profits. Each agent summarizes its own behavior in every τ time steps, calculates the revenues, and adjusts its skills in order to get more profits in the future.

A basic intuition has been widely believed that the tendency to implement an action should be reinforced if it produces favorable results and weakened if it produces unfavorable results [19]. A capability *biased learning mechanism* is introduced [2]. According to this mechanism, the agent will strengthen the capability resulting in more profits and weaken the capability resulting in fewer profits.

We define the increment of searching/exploiting efficiency, $\Delta K_s/\Delta K_e$, as the change of searching/exploiting efficiency in time interval τ.

$$\begin{array}{l} \Delta K_s = (k_1 + k_2 \cdot sign\,(\Delta W_s - \Delta W_e)) \cdot K_s \cdot (1 - K_s) \\ \Delta K_e = (k_1 + k_2 \cdot sign\,(\Delta W_e - \Delta W_s)) \cdot K_e \cdot (1 - K_e) \end{array} \tag{7}$$

Where $k_2 \geq k_1 \geq 0$ and $sign()$ is a sign function.

According to the definition, we can calculate the exploiting capability and the radius in the next time step, $E = (K_e + \Delta K_e) \cdot E_{max}$ and $r = \frac{\sqrt{log_{1-p}(1-K_s-\Delta K_s)}-1}{2}$.

2.4 Willingness Price and Market

In the market, we take into account the opportunity cost. Firstly we have to introduce the concept of the expected average productivity, which is defined as the productivity that an agent expects to work with in the next time step. It is calculated as $\hat{R} = R/(t_f + R/E)$, where, t_f is the average time to search a resource during the last τ time steps and R/E is the time to exploit the resource found. A buyer will have to spend R/E if it manages to get the pack of resources through purchasing, and a seller will save the exploiting time R/E if it successfully sells the information about the resources. Supposed the agent's expected average productivity $\hat{R}$ in the next time step and the market price of P, the former will finally get the revenue $R - \hat{R} \cdot R/E$, and the latter will obtain the profit of $P + \hat{R} \cdot R/E$.

However, Transaction cost has not been taking into account in the above section. In real life, trade always has costs that maybe incurred by searching who it is that one wishes to deal with, conducting negotiations leading up to a bargain, drawing up the contract, undertaking the inspection needed to make sure that the terms of the contract are being observed, and so on. Hence, in this model transaction cost is introduced. If a buyer pays the price of P, it will also pay a tax, $\alpha \cdot p, 0 \leq \alpha \leq 1$. In other words, a buyer has to pay $p \cdot (1 + \alpha)$.

As for the buyer, it will purchase the new information about resources under the condition that expected revenue to buy is higher than that of autarky,

$$R - \hat{R} \cdot R/E > p \cdot (1 + \alpha)\,. \tag{8}$$

As for the seller, it may trade the information about resources under the condition that expected revenue to sell is higher that that of autarky,

$$p + \hat{R} \cdot R/E > R. \tag{9}$$

From the above two conditions, we can make a conclusion that a buyer may pay no more than P_p and a seller may sell the information with the price that is no less than P_s,

$$\begin{aligned} P_p &= \left(R - \hat{R} \cdot R/E\right) / (1 + \alpha) \\ P_s &= R - \hat{R} \cdot R/E \end{aligned} \tag{10}$$

In fact, people often make decisions based on their previous experience, and recent experience generally plays a larger role than past experience in determining behavior [6]. Considering the impact of recent successful trade price, we write the final willing price as

$$\begin{aligned} P_p^* &= \tfrac{K_e}{K_e+K_s} \cdot P_p + \left(1 - \tfrac{K_e}{K_e+K_s}\right) \cdot P' \\ P_s^* &= \tfrac{K_e}{K_e+K_s} \cdot P_s + \left(1 - \tfrac{K_e}{K_e+K_s}\right) \cdot P' \end{aligned} \tag{11}$$

where P_p^* is a buyer's purchasing willingness price,P_s^* is a seller's selling willingness price and P' is the last *successful trade price*. At the beginning of evolutionary process, all agents' P' value are zero.

A successful trade always tends to involve the fact that the willingness price of the buyer isn't lower than that of the seller. In this model, an existing broker collects the information about the willingness prices of sellers and buyers which come into the market. Supposing that there are n_1 sellers and n_2 buyers in the market, we may list the series of selling willingness prices $P_{si}^*, i = 1, 2, \ldots, n_1$, in ascending order and the series of purchasing willingness prices $P_{pj}^*, j = 1, 2, \ldots, n_2$, in descending order. The dealer manages to find the maximum k that satisfies $P_{sk}^* \leq P_{pk}^*, k \leq n_1, k \leq n_2$. At last, the dealer decides the k sellers and the k buyers to make successful trades in the current time step. One of the k sellers matches one of the k buyers, and the trading price of P is the average value of their willingness prices, $\frac{P_p^* + P_s^*}{2}$.

3 Simulation Results

We set parameters as follows. $L = 3163, R = 200, \rho = 0.001, M = 100, \tau = 20, \alpha = 0.1, k_1 = 0.025, k_2 = 0.025, c = 0.005, E_{max} = 200$. Initial value for K_e and K_s are 0.1. Termination criterion is set as running time step 3000. We write the code in Java and Swarm, The programs are run on a Pentium 4 PC computer. A typical run of this computation takes 4 hours.

To track the evolving process of the model system, we employ several macro variables, such as total productivity, count of agents of different types, price distribution and wealth distribution. The evolving results of each macro variable are shown in following figures.

3.1 Total Productivity

Total productivity is defined here as the sum of incomes of all agents and total tax that happens at the same time, of which evolving process is shown in Fig. 2. To describe the evolving process of total productivity, we compare it with trade volume and non-seller count of which evolving process also presented in Fig. 2. Non-seller count at time step t is the count of sellers who fail to complete a trade at the previous time step $t-1$, is calculated by subtracting trade volume from the count of sellers at the time step $t-1$. In our model, we can take the trade volume as a measure for the extent of frequency of cooperation between agents. Figure 2 shows us that there exists similar increasing trend between trade volume and total productivity. This phenomenon conforms to the economic fact that specialization and cooperation can improve social productivity. The leaps in Fig. 2 and following figures are obviously observed and the cause will be presented in our further coming paper.

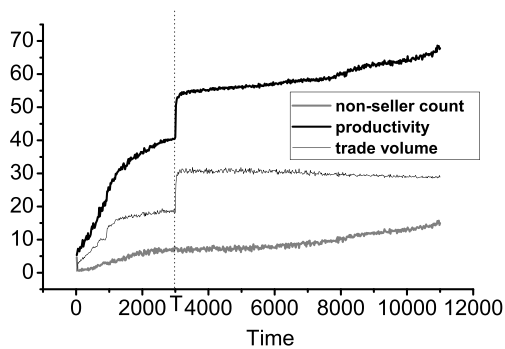

Fig. 2. x-coordinate is the time of evolving process and y-coordinate is the value. Curve of productivity in the pedal represents the evolving process of total productivity, and curve of non-seller count and trade volume represent respectively its own evolving process.

3.2 Market Forming

Market in our model is the indispensable media that helps to realize the cooperation among agents. The evolving process of market structure, which includes count of agents of different types and the willingness price distribution, are presented in Fig. 3. With time evolving, count of agents of different types changes, with leap phenomenon mentioned before, as shown in pedal B of Fig. 3. Overall, the count of agents in autarky decreases and the count of miners and searchers both increase, indicating that the extent of specialization and cooperation are intensified. The coupled phenomenon in the evolving process of willingness price distributions can be seen from pedal B in Fig. 3., that is, when time evolves forward, both distribution ranges of willingness prices to sell and to buy become

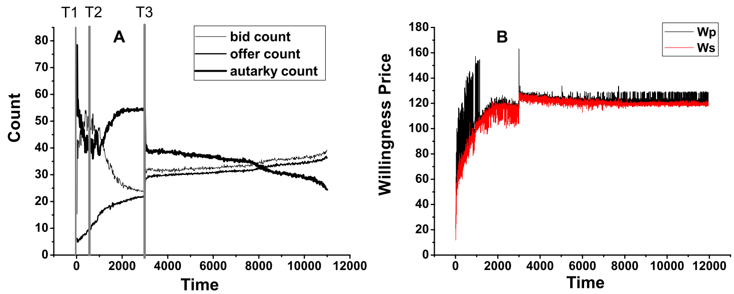

Fig. 3. Curves in Pedal A represent respectively the evolving process of the count of miners, of the count of searcher and the count of agents in autarky, according to curve signs. In pedal B, the upper line section Wp corresponding to each time describes the distribution range of willingness price to sell and the lower line section Ws describes the distribution range of willingness price to buy.

narrow and stable, with a increase trend in the beginning and stable oscillation at higher price later, indicating that the market need some time to shape, and can operates stably finally in our model. It may conform to fact of market evolving to some extent.

3.3 Wealth Distribution

Wealth distribution is also a very important variable in the evolving process of whole system. As shown in the Fig. 4A, most agents' wealth $W(t)$ increases while the system evolves except for only one "poor" agent that does not accumulate any wealth and goes negative. This agent is an outlier that, unfortunately, is initially located in some lean soil and nearly never finds or exploits one pack of resources. The daily consumption makes its wealth go to negative.

At the beginning, each agent has the same initial wealth value and the difference between the most rich and poor increases with time. Some agents have plenty of wealth, while others have less than that of the former.

If we do statistics on how many agents there are between one wealth value interval, we get frequency counts, as shown in Fig. 2B. If we take into account data points at only one time step in Fig. 2A, we can get only 100 data points and that is not statistical significant. So, the data points in Fig. 2B are got by this way. We consider all the data from time step 1 to time step 3000. In the log-log plot, the wealth frequency count demonstrates a local power law distribution. We do linear regression to the simulated data and get an exponent 0.44 for the first part and an exponent 10.02 for the latter part. The local power law distribution shows us that the wealth distribution before all the agents' wealth accumulation reaches 60000 is very "flat" and that after 60000 is very "steep". This means that this model showed us a society that enlarges the gap between the rich and the poor after the economy reaches certain threshold.

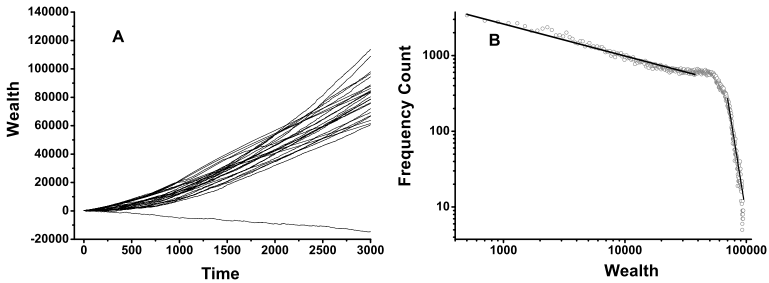

Fig. 4. A is the figure for each agent's wealth vs. time. There are 100 lines, which represents 100 agent's wealth. There is only one "poor" agent that goes to negative due to initially located in some lean soil and finds/exploits no resources. The wealth difference between the rich and poor widens vs. time even though all agents have the same initial wealth value. B is the figure for wealth frequency count. The agents' wealth is ranked in ascending order. The number of agents that have certain wealth values demonstrates a local power law distribution. Gray circles are the simulated wealth value counts and the black lines are linear regressions of the simulated results. The first part has an exponent 0.44 and the latter part has an exponent 10.02.

4 Conclusion

In this paper, we introduced an agent-based cooperation system in which agents can learn and adapt. Cooperation is a highlight in this paper, where agents collaborate in the sense that they trade information about resources and both buyers and sellers get benefit from the trade. Our learning mechanism belongs to individual learning [10], which is a biased learning mechanism, that contributed a lot to the division of labor.

The expectation mechanism is a novel and unique pricing mechanism that we do not find in any other previous papers [1, 3, 4, 7–18]. This may provide an agent a way of "real intelligence" that the agent can make decisions by "seeing in the future".

With the learning and pricing mechanism, our system evolves with some collective behaviors on specialization and cooperation, market forming and power-law wealth distribution, which is a commonly highlighted area in the literature [5, 20]. The correlation between the results of this model and stylized facts in the literature will be a focus of our next work. Our results verify once again the importance and convenience of agent-based modeling to the study on social complex systems. We can take more realistic conditions into account and make our model be more closer to real economy and real wealth distribution.

References

[1] Andriani, P.: Diversity, knowledge and complexity theory: some introductory issues. Int. J. Innovat. Manag. 5(2), 257–274 (2001)

[2] Baant, Z., Osman, E.: Double power law for basic creep of concrete. Mater. Struct. 9, 3–119 (1979)

[3] Bunn, D., Oliveira, F.: Agent-based simulation-an application to the new electricity trading arrangements of England and Wales. IEEE Trans. Evol. Comput. 5(5), 493–503 (2001)
[4] Conte, R.: Agent-based modeling for understanding social intelligence. PNAS 99, 7189–7190 (2002)
[5] Ding, N., Wang, Y., Xu, J., Xi, N.: Power law distribution in circulating money: Effect of preferential behavior. Int. J. Mod. Phys. B 18, 17–19 (2004)
[6] Erev, I., Roth, A.: Predicting how people play games with unique, mixed-strategy equilibria. Am. Econ. Rev. 88, 848–881 (1998)
[7] Hagg, T., Huberman, B.: Controlling chaos in distributed systems. IEEE Trans. Syst. Man Cyber. 21, 1325–1332 (1991)
[8] Holldan, J.: Hidden Order-how adaptation builds complexity. Addison-Wesley, Reading (1995)
[9] Izumi, K., Ueda, K.: Phase transition in a foreign exchange market-analysis based on an artificial market. IEEE Trans. Evol. Comput. 5(5), 456–470 (2001)
[10] Vriend, N.J.: An illustration of the essential difference between individual and social learning, and its consequences for computational tools. J. Econ. Dynam. Control 24, 1–19 (2000)
[11] Kephart, J., Hagg, T., Humberman, B.: Dynamics of computational ecosystems. Phys. Rev. A 40, 404–421 (1989)
[12] Kephart, J., Hagg, T., Humberman, B.: Collective behavior of predictive agents. J. Phys. D 106, 48–65 (1990)
[13] LeBaron, B.: Empirical regularities from interacting long- and short-memory investors in an agent-based stock market. IEEE Trans. Evol. Comput. 5(5), 442–455 (2001)
[14] Leigh, T.: Agent -based modeling of evolutionary economic systems. IEEE Trans. Evol. Comput. 5(5), 437–441 (2001)
[15] Lempert, R.: Agent-based modeling as organizational and public policy simulators. PNAS 99, 7195–7196 (2002)
[16] Newman, M.: Power laws, Pareto distributions and Zipf's law. Contemp. Phys. 46, 323–351 (2005)
[17] Savageb, M., Askenazi, M.: A swarm-based multi-agent ecological disturbance model technical report. Tech. Rep. STF-TR-98-06-056, Santa Fe Institute (1998)
[18] Stanley, H.: Similarities anddifferences between physisc and economic. J. Phys. A 299, 1–15 (2001)
[19] Sutton, R., Barto, A. (eds.): Reinforcement Learning: An Introduction. Cambridge University Press, Cambridge (1998)
[20] Xie, Y., Wang, B., Hu, B., Zhou, T.: Power law distribution of wealth in population based on a modified EZ model. Phys. Rev. E 71, 046, 135(2005)