MEASURING HETEROGENEITY IN STOCHASTIC MODELS

Florentin PALADI, Denis DOŢENCO, Visarion GAMURARI

Catedra Fizica Teoretică

Un model teoretic care descrie interacţiunea într-o populaţie eterogenă de agenţi a fost dezvoltat şi analizat în cadrul discuţiilor recente despre decalajul dintre modelele computaţionale ABM şi modelele analitice stocastice [1]. A fost demonstrat că acest model probabilistic descrie exact rezultatul previzibil al simulărilor computaţionale care trebuie să fie suficient de lungi pentru a echilibra starea sistemului. Acest rezultat depinde de distribuţia stărilor în sistem şi este definit atât cu ajutorul valorii medii, cât şi al coeficientului de variaţie pentru *payoff*. Ultima mărime este independentă de scara valorilor şi denotă cât de inegale sunt distribuţiile în sistem, iar, prin urmare, şi nivelul de eterogenitate al modelului nostru stocastic. Rezultatele obţinute sunt raportate pentru valoarea medie şi coeficientul de variaţie pentru *payoff*, calculate pentru modelul prezentat în [1]. În lucrare nu este specificată o formă funcţională particulară pentru a evita alegeri arbitrare, în cazul unui sistem format din patruzeci agenţi cu 1221759 distribuţii posibile în şase clustere după trei stări diferite.

Introduction

We have recently introduced a theoretical model which describes the interactions in a heterogeneous population of agents [1]. In particular, there are *N* entities which can be in 3 different states (call them *Left, Center* and *Right*), and can play 3 actions (again *Left, Center* and *Right*). Interaction in this agent-based model always involves one *active* and one *passive* player, but the agents can play both roles interchangeably. They have preferences over their states: love one state, are neutral with respect to another state and they hate the remaining state. If the active player follows the *first rule*, it always plays the action corresponding to the loved state. If it follows the *second rule*, it randomises between actions corresponding to the loved and neutral states. When two agents meet, the active player sets the passive player's state according to his action, which in turn is determined by one of the applied rule. This identifies only 6 possible combinations. Denote with probabilities $p_1 ... p_6$ the shares of the population characterized by each combination of preferences. That is, drawing randomly one agent, it will be of type *i* with probability p_i. After each interaction, the passive player gets a payoff of +1 if it is in the loved state, a payoff of 0 if it is in the neutral state, and a payoff of –1 if it is in the hated state. The active player does not get any feedback. Let $N = 1, 2, \dots, \infty$ be the total number of entities in the model, and $\{n_1, n_2, n_3, n_4, n_5, n_6\}$ is their partition into *m*=6 subsets. Each subset can be called *cluster*, and the process itself – *clustering*. The size of each cluster can vary from 0 to *N*, $n_i = \overline{0, N}, i = \overline{1, 6}$, and $\sum_{i=1}^{6} n_i = N$. The number of possible partitions *P* is a function of *N*, and the solution is $P(N) = \frac{1}{5!}\prod_{i=1}^{5}(N+i)$ [1].

The purpose of this paper is to report the results for the mean and the coefficient of variation of the payoffs for the model proposed in [1] in the case of a system formed by forty agents with *P*(*N*)=1221759 possible partitions into six clusters over three different states, and to prove that even a relatively simple stochastic model can describe precisely the expected outcomes from corresponding agent-based computational simulations [2, 3].

The model

Consider an active agent of type *1* (loves *Left* and hates *Center*) which meets in turn all other (passive) agents, including himself. If it follows the first rule, then it will play *Left* causing a payoff of +1 in $(p_1+p_2)N$ agents, and a payoff of –1 in $(p_3+p_5)N$ agents. Note that there are $(p_1+p_2)N$ similar entities in the ensemble. Suppose now that everybody meets everybody else both as active and as passive agent. So the average payoff when everybody plays according to the first rule is

$$\pi_1 = (p_1+p_2)(p_1+p_2-p_3-p_5)+(p_3+p_4)(-p_1+p_3+p_4-p_6)+(p_5+p_6)(-p_2-p_4+p_5+p_6). \quad (1)$$

Similarly, the average payoff with the second rule is

$$\pi_2 = \frac{1}{2}(p_1(p_1-p_3-p_5+p_6)+p_2(p_2-p_3+p_4-p_5)+p_3(-p_1+p_3+p_5-p_6) + p_4(-p_1+p_2+p_4-p_6)+p_5(-p_2+p_3-p_4+p_5)+p_6(p_1-p_2-p_4+p_6)). \quad (2)$$

To study the behaviour of $\pi_1 - \pi_2$ we used to set some of the probabilities to zero [1]. In this paper we will not specify a particular functional form in order to avoid arbitrary choices: we represent the distribution of states as a single point in a three dimensional space, where the axes are labelled *l*, *c* and *r*. The *l* coordinate is found by counting all agents who love *Left*, and subtracting all agents who hate *Left*. The result is then normalized to the size of the population. Similarly for the other two coordinates. Hence,

$$\begin{aligned} l &= p_1 + p_2 - p_3 - p_5 \\ c &= p_3 + p_4 - p_1 - p_6 \\ r &= p_5 + p_6 - p_2 - p_4 \end{aligned} \tag{3}$$

and $l + c + r = 0$.

Note that different distributions of states can lead to the same point in the sphere. For instance, the point in the origin is given not only by $p_1 = p_2 = \ldots = p_6 = 1/6$, but by any combination of preferences such as $p_1 = p_3$, $p_2 = p_5$, and $p_4 = p_6$.

We can now define the polarization of states as the distance from the center of the sphere:

$$d(l,r,c) \equiv d(p_1, p_2, \ldots, p_6) = \sqrt{l^2 + r^2 + c^2}\ . \tag{4}$$

Note that $d \in \left[0, \sqrt{2}\right]$: all points thus lie inside a sphere around the origin.

The variances $\sigma^2{}_1$ and $\sigma^2{}_2$ are defined for each discrete distribution $D \equiv 1, 2$ with the expectation (mean) value π_D as follows:

$$\sigma_D^2 = \sum_{i=1}^{6} p_i (\pi_{i,D} - \pi_D)^2\ , \tag{5}$$

where

$$\begin{aligned} \pi_{1,1} &= p_1 + p_2 - p_3 - p_4, & \pi_{4,1} &= p_3 + p_4 - p_5 - p_6 \\ \pi_{2,1} &= p_1 + p_2 - p_5 - p_6, & \pi_{5,1} &= -p_1 - p_2 + p_5 + p_6 \\ \pi_{3,1} &= -p_1 - p_2 + p_3 + p_4, & \pi_{6,1} &= -p_3 - p_4 + p_5 + p_6 \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} \pi_{1,2} &= (p_1 - p_3 - p_5 + p_6)/2, & \pi_{4,2} &= (-p_1 + p_2 + p_4 - p_6)/2 \\ \pi_{2,2} &= (p_2 - p_3 + p_4 - p_5)/2, & \pi_{5,2} &= (-p_2 + p_3 - p_4 + p_5)/2 \\ \pi_{3,2} &= (-p_1 + p_3 + p_5 - p_6)/2, & \pi_{6,2} &= (p_1 - p_2 - p_4 + p_6)/2 \end{aligned} \tag{5.2}$$

Since variance is scale-sensitive, it makes little sense to use it as a measure of dispersion when the mean values can significantly differ. We, thus, should divide the standard deviation by the mean to obtain the coefficient of variation, which is scale-free.

Results

Figure 1 explores how the outcome varies as a function of the distance *d*. The whole range [0,1] is sampled, for all probabilities $p_1 \ldots p_6$. The step considered for creating all combinations of probabilities is 0.025, i.e. the total number of agents is 40. The average values for π_1 and π_2 are shown in Figure 1a, and for each value of the distance from the center of the sphere, $d(l,r,c)$, the frequency of wins with each rule is computed (Figure 1b). When $\pi_1 - \pi_2 > 0$ a win is assigned to the first rule, and when $\pi_1 - \pi_2 < 0$ a win is assigned to the second one. Exactly in the center of the sphere the two rules lead to the same payoff, independently of the underlying distribution of states. Close to the center, each rule wins in about 50% of the cases. Then, as we move away from the center, the first rule improves its performance, and is always better when the states are totally polarized. Meanwhile, the total number of states for intermediate values of the distance *d* is much larger than for the dispersed and polarized states.

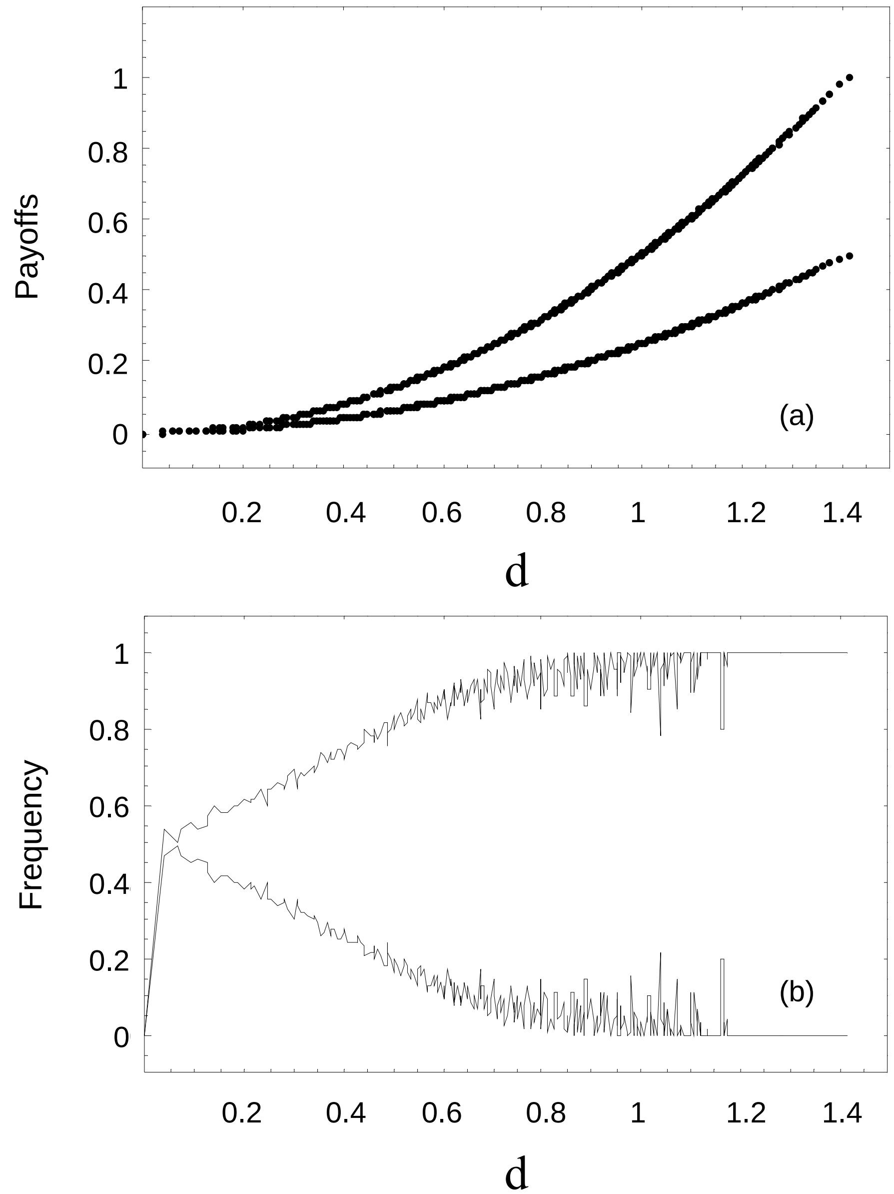

Fig.1. Average values for π_1 (upper line) and π_2 (a), and relative frequency of negative and positive values of $\pi_1 - \pi_2$ (b).

Figure 2 shows that for any level of fragmentation of states, the two rules give the same average values of the coefficient of variation on the distance *d*. An interesting result was also reported [2]: on average, however, when one rule is better in terms of higher expected payoffs it is also better in terms of lower heterogeneity.

Conclusions

The model developed in the paper [1] was used to study agent-based interactions in heterogeneous systems. We can conclude that, depending on the underlying distribution of the states, both rules can be optimal. However, as the states become more polarized, i.e. for a less heterogeneous system, the first rule clearly takes the lead. For a homogeneous system, i.e. the same preferences, both variances σ^2_1 and σ^2_2 go to zero. In an agent-based computational model, in the case of the first rule, everybody plays the same action and gets the same payoff (+1$\cdot$ *N*, where *N* is the population size). In the case of second rule, two actions can be played, causing either a payoff of +1 or a payoff of 0 to the passive player, at each interaction. However, since everybody is playing against everybody else, as the system size gets larger everybody finally gets a payoff of 0.5$\cdot$*N*.

However, the variance in the first case is generally higher than the second one, especially when the states are dispersed. Although in general different, the two scale-free coefficients of variations give exactly the same value when they depend on the distance *d*.

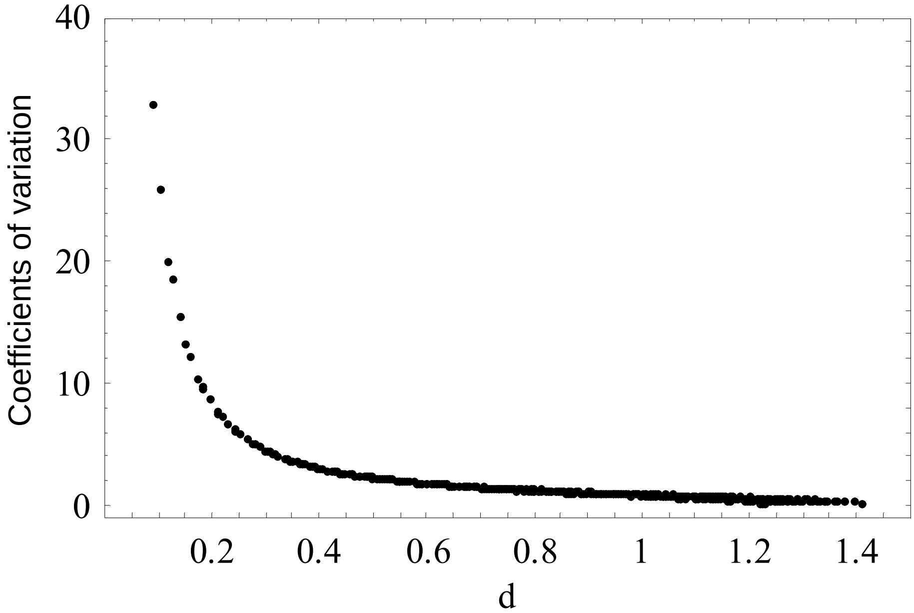

Fig.2. Average values for the coefficients of variation.

Finally, it is worth to mention that the model is a general one and allows simulations for any size of the system. We came here to the choice of forty entities just due to the restriction in computer memory: there are over one million values (1221759) for each quantity described in this paper!

References:

1. Paladi F. Bridging the gap between stochastic and agent-based computational models // Analele Ştiinţifice ale Universităţii de Stat din Moldova. Seria „Ştiinţe fizico-matematice". - Chişinău: CEP USM, 2005. - P.78-81.
2. Richiardi M. From moral to social norms and back // Working paper. - 2006.
3. Paladi F. Sisteme complexe: modele analitice şi numerice avansate de tip cluster. - Chişinău: CEP USM, 2005. - 174 p.

Acknowledgements
Authors wish to thank Prof. A.Barsuc and Dr. M.Richiardi for useful suggestions.
Financial support from institutional research grant 06.408.036-F is gratefully acknowledged.

Prezentat la 26.01.2007