A MODEL OF MANY GOAL-ORIENTED STOCHASTIC AUTOMATA WITH APPLICATION ON A MARKETING PROBLEM

by

Y.M. EL-FATTAH

University of Liège

Liège, Belgium

ABSTRACT

Here presented is a model of many goal-oriented stochastic automata. The goal of each automaton is the extremum of the absolute mean value of a certain utility function. That function depends explicitly on the automaton strategy and the environment response. Without need of any a priori knowledge each automaton adapts its structure in the process of achieving its own goal. By suitably setting the environment characteristics the automata model can be useful for the analysis of some operations research problems. As example that model is used for the study of the price formation process in a free competitive market. The results demonstrated the convergence of the automata updating scheme as well as the influence of a number of interesting physical parameters (like the buyers tactics, the sellers psychology, etc...) on the equilibrium condition.

1. INTRODUCTION

Tsetlin [1] has proposed different norms of behavior of a finite automaton working in a random environment. In that work the environment is assumed to either penalize or reward each action of the automaton according to certain unknown probabilities. The bahavior of an automaton is called expedient if the average penalty is less than the value corresponding to choosing all actions with equal probabilities. The behavior is called optimal or ε- optimal according to whether the average penalty is equal or arbitrarily close, respectively, to the minimum value. Krylov and Tsetlin [2] introduced the concept of games between automata and studied in particular Two-Automaton Zero-Sum games.

Stochastic automata with variable structure have been introduced by Varshavskii and Vorntsova [3] to represent learning automata attemptting a certain norm of behavior in an unknown random environment. Since the date of that work a respectable number of works has appeared, studying different aspects of learning automata and applying it in simulating very simple norms of behavior (like that introduced by Tsetlin) and also simple automata games (such as Two-Automaton Zero-Sum games). For a survey on the subject we refer to Narendra and Thathacher [4].

The contribution of this paper is to direct the attention of using learning automata to simulate an important class of problems of collective behavior whose deterministic version has been the subject of recent investigation mainly by Malishevskii and Tenisberg, see [5] - [7]. In that class of problems there exists a type of relation in the collective where the behavior of the participants possesses a definite mutual opposition. Such situation can arise for example in economic systems : the case of price regulation in a competitive market [8]; or in management systems : the problems of resource allocation [9].

In the model introduced in this paper a collectice of interacting stochastic automata is considered. Each automaton has a behavioral tactic directed towards the realization of its own goal, taken to be the extremun of a certain utility function. That function depends explicitly on the automaton strategy and the environment response. The automata interactions arise from the dependence of the environment response on the whole set of strategies used by the collective of automata. That dependence is generally stochastic and unknown to all the automata. Furthermore, any automaton does not know neither the utility functions, nor even the number of the other automata. The only available knowledge to each automaton is the realization of its utility function following the use of a certain strategy.

The use of automata game to model the process of market price regulation (or optimization), described in Karlin [8], from the viewpoint of collective behaviour was demonstrated by Tenisberg [5]. In his work, Tenisberg [5] made two assumptions. The first assumption is connected with the substitution of the probability characteristics of the buyers' demand by deterministic characteristics (mean value). This is equivalent to the assumption that the transactions are sufficiently numerous. The second assumption is that already in a time small compared to the characteristic time of the system a fairly large number of interactions between the sellers and the buyers occur. These assumptions permit the model to be described approximately by deterministic differential equations. Later Malishevskii and Tenisberg [6] formulated theorems about the existence and uniqueness of the equilibrium situation in the game (in the sense of Nash) and the attainability of this situation in the process of the automata game.

Krylatykh [10] modified the deterministic model of Tenisberg [5] by taking into consideration the psychological attitudes of sellers and buyers with respect to the market situation. This has the effect of including nonlinear utility function instead of a linear one in the Tenisberg model [5]. This corresponds to the cases when the automata are not able to perceive the created situations adequately; simulating the individual psychological peculiarities of the buyers and sellers, who in reality are not necessarily always objective (corresponding to a linear utility function).

Applying the present stochastic automata model the assumptions of Tenisberg [5] are not needed; in particular the stochastic nature of the buyers demand will be respected. In addition the stochastic model reveals the effect of a number of interesting factors - like the sellers psychology - on the modes of collective behaviour which has no analog in deterministic modelling as shown in section 5.

2. AUTOMATA MODEL

As model of collective behavior we consider the following game of N stochastic automata $A^1, A^2, \ldots, A^N$. The automata operate on a discrete time scale $t = 0,1,2,\ldots$ The input $s^i(t)$ to each i-th automaton can acquire one of the values $s_1, s_2, \ldots, s_m$. The output $f^i(t)$ of the automaton A^i will be assumed to take one of the k_i values $f_1^{\,i}, \ldots, f_{k_i}^{\,i}$ which will be called its strategies. We will say that the automaton A^i uses the j-th strategy if $f^i(t) = f_j^{\,i}$.

A play f(t) carried out at time t will be the name given to a set $f(t) = (f^1(t), \ldots, f^N(t))$ of the strategies used by the automata $A^1, \ldots, A^N$ at time t. The outcome s(t) of a play f(t) is a set $s(t) = (s^1(t), \ldots, s^N(t))$ of the refree or environment responses at time t. The model is depicted schematically in Fig.1. The environment is completely characterized by the probability $P(f(t), s(t))$ of the outcome s(t) for every play f(t). As only stationary environments will be considered, the aforementioned probability can simply be written as $P(f,s)$. The game of N automata A^i is considered to be a game with independent outcomes, i.e.

$$P(f,s) = \prod_{i=1}^{N} P^i(f,s^i) \tag{2.1}$$

Let us introduce the indicator functions $\Phi^i(f)$ defined by

$$\begin{aligned}\Phi^i(f) &= M_s^i\{ F^i [\theta^i(f^i,s^i)]\} \\ &= \sum_s F^i[\theta^i(f^i,s^i)]P^i(f,s^i) \quad (i=1,\dots,N)\end{aligned} \tag{2.2}$$

where M_s^i denotes the mean value with respect to s^i, F^i is an utility function of the incentive (or utility) θ^i. The incentive θ^i depends explicitely on the automaton strategy f^i and the environment response s^i.

The objective of each i-th automaton is to choose its strategy f^i in order to minimize the absolute value of its own indicator function (2.2), i.e. to minimize

$$Q^i(f) = |\Phi^i(f)| \tag{2.3}$$

In classical N-person games, each player possesses an adequate a priori knowledge of the game, i.e. the criteria and the sets of pure strategies for all the players. It is defined that the i-th player uses the mixed strategy $p^i = (p_1^i,\dots,p_{ki}^i)$ if the uses his pure strategy f_j^i with probability $p_j^i(j=1,\dots,k_i)$, $\sum_{j=1}^{k_i} p_j^i=1$). Nash's basic theorem states that any finite N-person game has at least one equilibrium situation in mixed strategies $(p^1,\dots,p^N)$, that is

$$\bar{Q}^i(p^{1*},\dots,p^{i*},\dots,p^{N*}) \le \bar{Q}^i(p^{1*},\dots,p^i,\dots,p^{N*}) \tag{2.4}$$

for all p^i, $i=1,\dots,N$, where

$$\bar{Q}^i(p^1,\dots,p^N) = \sum_{j_1 j_2 \dots j_N} p_{j_1}^1 p_{j_2}^2 \dots p_{j_N}^N Q^i(f_{j_1}^1, f_{j_2}^2,\dots,f_{j_N}^N) \tag{2.5}$$

also denoted by $\bar{Q}^i(p)$, is the mathematical expectation of the gain of the i-th player when the set of mixed strategies $p=(p^1,\dots,p^N)$ is used.

Unlike N-person, players automata in automata games do not possess any a priori information about the game. They know neither about the criteria (2.3) nor even the number of game partners. They must choose their strategies (or which is the same their probability vectors p^i) in the course of the game by using the only available information : the realizations of their incentive functions $\theta^i(f^i,s^i)$. In studying automata games we thus come to know the behavior of the players in the game process.

The indicator functions $\Phi^i(f), i=1,...,N$ given by (2.2) are assumed to satisfy the conditions of individual and group contramonotonicity [6], i.e. for any subset I of the set of indexes $\{i\}=\{1,..,N\}$ the function.

$$\Phi^I = \sum_{i\varepsilon I} \Phi^i(f)$$

decreases in the set of own variables f^i, $i\varepsilon I$ and does not decrease in the set of foreign variables $f^j, j\varepsilon \bar{I}$ ($\bar{I}$ is the complement of I).

If each automaton A^i knows its own indicator function Φ^i, and its strategy f^i can take any value in the continuous interval $[f_1{}^i, f_{ki}{}^i]$ then the optimal tactic can be given by the simple differential relation [5] - [7]

$$\overset{\circ}{f}{}^i(t) = \bar{\Phi}^i(f(t)) \tag{2.6}$$

where

$$\bar{\Phi}^i(f) = \begin{cases} 0 & \text{if } f^i=f_1{}^i, \Phi^i < 0 \text{ or } f^i=f_{ki}{}^i, \Phi^i > 0 \\ \Phi^i(f) & \text{in all other cases.} \end{cases} \tag{2.7}$$

This means that the trajectories of the system (2.6) converge to the Nash point f^* if exists.

Let us emphasize that in the present model of collective behavior the goal of each automaton is not fully determinate, i.e. known only up to certain parameters for which there is no a priori information. Specifically automaton A^i does not know its own indicator function $\Phi^i(f)$; all what is known to it is the incentive function $\theta^i(f^i,s^i)$, see eqn. (2.2).

Let us arrange the set of strategies of the i-th automaton such that $f_j^i > f_k^i$ for all $j > k$. For any strategy $f_j^i (1 < j < k_i)$ we call f_{j+1}^i the next supremal strategy and f_{j-1}^i the next infimal strategy. The strategies f_1^i and f_{ki}^i are the infimal and supremal absorbing strategies, respectively of the i-th automaton.

Inspired by the behavioral conception represented by eqn.(2.6) we propose the following model of learning automata.

Let us introduce the functions,

$$u^i = \begin{cases} +1 & \text{if } \theta^i > 0 \\ 0 & \text{if } \theta^i = 0 \\ -1 & \text{if } \theta^i < 0 \end{cases} \tag{2.8}$$

and

$$\bar{\theta}^i(f^i,s^i) = \begin{cases} 0 & \text{if } \quad f^i=f_1^i, \theta^i < 0, \text{ or } f^i=f_{k_i}^i, \theta^i > 0 \\ \theta^i(f^i,s^i) & \text{in all other cases} \end{cases} \tag{2.9}$$

The idea underlying the functioning of a learning automaton in the present model can be loosely stated as follows. At any time step provided that an automaton is not at either of the absorbing states and the automaton action has elicited an environment response for which the incentive function θ^i is greater than zero than at the next time step the probability of the next supremal action is increased; on the other hand if the incentive function is less than zero then the probability of the next infimal action is increased. Otherwise, that is if the automaton is at either of the absorbing states or the incentive function is zero, the automaton remains in the status quo. That idea can be analytically represented by the following updating scheme for the automata strategis. Provided that $f^i(t) = f_j^i$ then

$$\begin{aligned} p^i_{j+u^i}(t+1) &= p^i_{j+u^i}(t) + \gamma(t+1)u^i F^i(\bar{\theta}^i(f^i(t),s^i(t))) \\ p^i_m(t+1) &= p^i_m(t) - \frac{\gamma(t+1)}{N+1} u^i F^i(\bar{\theta}^i(f^i(t),s^i(t))), \\ & \qquad m=1,..,k_i, m \neq j + u^i \end{aligned} \tag{2.10}$$

where $\gamma(t)$ satisfies the classical conditions of stochastic approximation schemes

$$\gamma(t) > 0 \quad , \quad \sum_{t=1}^{\infty} \gamma(t) = \infty \quad , \quad \sum_{t=1}^{\infty} \gamma^2(t) < \infty \tag{2.11}$$

A function block diagram for each i-th automaton may be represented as shown in Fig. 2.

3. ENVIRONMENT MODEL

As said before the environment is completely characterized by the probability $P(f,s)$ of the outcome $s(t)$ for every play $f(t)$. That probability also fully specifies the interaction between the automata.

In the following we present two different models of the environment, named the "pairwise comparison" and the "proportional utility".

3.1. Pairwise comparison.

Let the environment be constituted of υ elements $j=1,..,\upsilon$. The j-th element finds out the strategies f^i and f^k of two randomly chosen (with equal probabilities) automaton, the i-th and the k-th($i,k=1,..,N$). The j-th element then responds in a probabilistic manner to only one of the chosen pair of automata; say with probability $p^j(f^i,f^k)$ to the i-th and with probability $p^j(f^k,f^i)=1-p^j(f^i,f^k)$ to the k-th.

We shall assume $p^j(f^i,f^k)=\psi(\delta^j(f^i,f^k)) = 1/2 + \mu(\delta^j(f^i,f^k))$ where $\delta^j(f^i,f^k)$ is a certain utility index for the j-th element, and $\mu(x)$ is a monotonically increasing odd function $\mu(+\infty)=\mu(-\infty) = 1/2$.

The total probability of a response from the j-th element of the environment to the i-th automaton can be written thus

$$p^{jni}(f) = \frac{2}{N(N-1)} \sum_{\substack{\ell=1 \\ \ell\neq i}}^{N} \psi(\delta^j(f^i,f^\ell)) \tag{3.1}$$

Notice that

$$0 \leq p^{j,i} \leq 1 \quad , \quad \sum_{i=1}^{N} p^{j,i} = 1 \tag{3.2}$$

The response of the j-th element of the environment to the i-th automaton is considered to be in the form

$$s^{j,i} = \omega^j(f^i) \tag{3.3}$$

where $\omega^j(.)$ is a piecewise continuous function.

Indeed eqs.(3.2) and (3.3) permit to write the conditional probability of the environment response s^i for a play f of the automata as follows,

$$p^i(s^i \leq \sum_{j=1}^{\ell} s^{j,i}/f) = \sum_{j=1}^{\ell} p^{j,i}(f) \quad , \quad (\ell=1,..,\upsilon) \qquad (3.4)$$

3.2. Proportional utility.

In this model each element of the environment responds to the automata with probabilities proportionable to the utilities of their strategies. The probability of a response from an element increases as the utility of an automaton strategy increases and becomes maximum for maximum utility. Hence the probability that the j-th element responds to the i-th automaton can be expressed thus,

$$p^{j,i}(f) = \phi(\delta^j(f^i))/\sum_{j=1}^{N}\phi(\delta^j(f^i)), j=1,..,\upsilon; \quad i=1,..,N \quad (3.5)$$

where $\phi(.)$ is some positive non-decreasing function, and $\delta^j(f^i)$ is the utility of the i-th automaton strategy f^i for the j-th element.

Eqs.(3.3) and (3.4) again complete the environment model description after replacing the probabilities $p^{j,i}(f)$ in eqn.(3.4) by the expression of eqn.(3.5).

4. APPLICATION - Price Formation in a Competitive Market.

Consider N sellers in a market trading in one specific commodity. Each i-th seller (i=1,..,N) is assumed to be supplied by a constant q^i units of that commodity per time increment (the interval between any two successive time steps). The strategy of any i-th seller f^i represents the price he specifies for his commodity. Let the i-th seller receives a demand π^i in monetary units for buying his commoditu at the specified price f^i. The financial incentive for the i-th seller is simply the difference between the demand and supply in monetary units, i.e.

$$\theta^i = \pi^i - q^i f^i \quad , \quad (i=1,..,N) \qquad (4.7)$$

The utility of that incentive may be interpreted differently by the sellers; each according to his psychological type. That interpretation is embodied in the utility function $F^i(.)$ of an i-th seller which may be considered in the following form,

$$F^i(\theta^i) = a^i(\exp(b^i\theta^i)-1) + d^i\theta^i \quad , \quad (i=1,..,N) \qquad (4.8)$$

The constants a^i, b^i, and d^i simulate the psychological type of the i-th seller as follows,

$$\begin{array}{lll} \text{Cautious type} & : a^i, b^i < 0 \;, & d^i = 0 \\ \text{Objective type} & : a^i, b^i = 0 \;, & d^i > 0 \\ \text{Hazardous type} & : a^i, b^i > 0 \;, & d^i = 0 \end{array} \qquad (4.9)$$

The nonlinearity of the utility function F^i for a cautious or hazardous seller indicates the lack of objectivity of such psychological types. Thus a hazardous type overestimates the importance of a good deal ($\delta^j > 0$) and underestimates the importance of the deal in the opposite situation ($\delta^j < 0$). A cautious type overestimates the importance of bad deals and underestimates the good ones.

The objective of each seller is to find a price strategy maximizes its utility function, or what amounts to the same ensures the least harmful situation (according to a certain psychology) created by the mismatch between commodity supply and demand in monetary units. Hence, each seller attempts to minimize the function (2.3) where the indicator function Φ^i is given by eqs. (2.2),(4.7), and (4.8).

The automata scheme (2.10) is used to simulate the behavior of the sellers.

In the case of pairwise comparison tactic of the buyers [5] the environment is simulated as in sec. 3.1. In this case the utility of the j-th buyer making his purshase from the i-th seller is given by

$$\delta^j(f^i, f^k) = f^k - f^i \qquad (4.10)$$

The ψ function in eqn. (3.1) may be taken thus,

$$\psi(x) = \begin{cases} 1 & x > \Delta \\ (x+\Delta)/2 \quad , & -\Delta < x < \Delta \quad , \\ 0 & x < -\Delta \end{cases} \qquad (4.11)$$

Here $(-\Delta,\Delta)$ represents the "active zone of the function". The function ω^j in eqn. (3.3) may be considered as

$$\omega^j(f^i) = \begin{cases} \beta^j & , \beta^j \geq f^i \\ 0 & , \beta^j < f^i \end{cases} \qquad (4.12)$$

This means that the purshasing transaction between the j-th buyer and the i-th seller will be completed only when the buyer's available amount of money β^j equals or exceeds the price f^i of a unit of the commodity.

In the case of reference prices tactic of the buyers [5] the environment is simulated as in section 3.2. In this case the utility of the j-th buyer making his purchase from the i-th seller is given by

$$\delta^j(f^i) = h^j - f^i \qquad (4.13)$$

where h^j is the reference price of the j-th buyer. The function $\phi(.)$ is taken to be the same as $\psi(.)$ defined by (4.11).

5. SIMULATION RESULTS.

In all the simulation experiments the following market parameters are considered,

Number of sellers $N=3$, Number of buyers $\upsilon=12$

Sellers' psychology :
Cautious $a^i=-1$, $b^i=-0.005$
Objective $d^i=0.02$
Hazardous $a^i=1$, $b^i=0.005$

Commodity supply $q^1=2$, $q^2=2$, $q^3=3$
Available money to each buyer $\beta^j=150$
Buyers' utility $\alpha=0.05$, see eqs. (4.10),(4.13)
The sequence $\gamma(t)$, see eqn. (2.10), were taken as

$$\gamma(t) = \frac{\gamma_o}{t} \quad , \quad \gamma_o=\text{const.}, \ t=1,2,.. \qquad (5.1)$$

The sellers sets of prices are first taken as :

c_k^i k \ i	1	2	3
1	100	100	140
2	140	130	170
3	-	170	200

The initial price probabilities for the different sellers are assumed

$p_k^i[o]$	$k \backslash i$	1	2	3
	1	0.5	0.2	0.5
	2	0.5	0.5	0.3
	3	-	0.3	0.2

5.1. "<u>Pairwise price comparison</u>" tactic.

Let the number of buyers $\upsilon=12$, and consider different psychological classes of sellers.

5.1.1. <u>Objective sellers</u>.

The width of the active zone, see eqn. (4.11) was taken first as $\Delta=200$. The effect of the constant γ_0, see eqn. (5.1), on the convergence of the sellers price probabilities was examined. It is concluded that a very small value of γ_0 (i.e. $0 < \gamma_0 << 1$) leads to a very sluggish convergence. On the other hand a value of γ_0 as big as 1 leads to a rather vigorous and oscillatory convergence. An optimum value for γ_0 seems to exist somehow in between, see Fig. 3.

Taking $\gamma_0 = 0.1$, a satisfactory convergence has been attained at n = 100. At that time step, the average of price probabilities (corresponding to 10 trials) are

<u>Table 1</u>

$E\{p_k^i[100]\}$	$k \backslash i$	1	2	3
	1	0.0528	0.0010	0.6708
	2	0.9472	0.4972	0.3292
	3	-	0.5018	0

which, presumably, is close to the equilibrium point.

The influence of the width Δ of the active zone was also tested with objective types of sellers and $\gamma_0 = 0.1$. With $\Delta = 2$ (which means that ψ was equal to 0 or 1 when the prices were different) we obtained the following mean probability matrix (the mean of 10 trials) :

Table 2

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.0355	0	0.5961
2	0.9645	0.3397	0.4039
3	-	0.6603	0

and with Δ = 2000 we obtained :

Table 3

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.1154	0.0010	0.6827
2	0.8846	0.5133	0.3173
	-	0.4857	0

The last two tables demonstrate the tendency to increase the prices as Δ gets smaller, which can be demonstrated as follows .

According to eqn. (4.11) of the function ψ , it is clear that if as a result of pairwise comparison

$$\max\{|\delta^j(f^i,f^k) \;,\; |\delta^j(f^k,f^i)|\} \geq \Delta \tag{5.2}$$

then the j-th buyer making that comparison will be definitely captured by one of them; specifically by the i-th if $\delta^j(c^i,c^k) \geq \Delta$ and by the k-th if $\delta^j(c^k,c^i) \geq \Delta$. An uncertain decision by the j-th buyer only happen when the prices are fairly close so that $\max\{|\delta^j(c^i,c^k)|,|\delta^j(c^k,c^i)|\} < \Delta$. In this case we shall say that "active competition" exists between the i-th and k-th sellers; we shall call the interval $(-\Delta,\Delta)$ the "active zone" of the function $\psi(x)$.

As $\Delta \to 0$, the competition by the prices tends to be uneffective; as the demand will be basically determined by the money flux into the market, and the commodity supply available to each seller. In such condition, it seems natural that each seller attempts to specify the highest possible price for his commodity. If the commodity supplies to the different sellers vary only slightly, the limit prices tend to be almost the same. This is verified by the simulation results.

Also, as in the deterministic case (cf. Tenisberg [5]), the effect of a small Δ, is more or less an equalization of prices in the market. This can be demonstrated by changing the set of prices of the second seller to include the price 140 instead of 130, as well as changing the number of buyers to have no abundant demands. Thus, take $\upsilon = 7$ and

c_k^i k \ i	1	2	3
1	100	100	140
2	140	140	170
3	-	170	200

With objective sellers and $\Delta = 2$ we obtained :

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.1674	0.0607	0.8670
2	0.8326	0.7464	0.1330
3	-	0.1929	0

This demonstrates that the sellers have all increased the probability of the price 140 which indicates the tendency of equalization of prices, compare with Table 5.

5.1.2. Hazardous sellers.

We put $\gamma_0 = 0.1$ and $\Delta = 200$. The mean probability matrix at $n = 100$ became

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.0457	0	0.3265
2	0.9543	0.1639	0.6735
3	-	0.8361	0

which, presumeably, is fairly close to the equilibrium point.

5.1.3. Cautious sellers.

We put $\gamma_0 = 0.1$ and $\Delta = 200$. The mean probability matrix at $n = 100$ became :

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.3715	0.0009	0.9332
2	0.6285	0.7919	0.0668
3	-	0.2072	0

which, according to the simulation trials, should be close to the equilibrium point.

Compared to the objective case, the last two tables demonstrate that the hazardous sellers tend to increase the probability of higher prices, while the cautious sellers tempt to increase the probability of lower prices. This result has no analog in deterministic modelling; where the sellers psychology does not affect the equilibrium prices, cf. Krylatykh [10]. In stochastic modelling, however, the expectation of the utility being zero does not imply that the expectation of the utility is zero due to the nonlinear from of the utility function in the case of hazardous or cautious types. In any case; the previous result agrees more with intuition and favors stochastic modelling for more realistic simulation. Let us now reduce the number of buyers. Take $\upsilon = 7$. This means less money flow into the market. All the other market parameters remain the same as before.

Considering objective sellers, the following mean price probabilities have been reached at n = 100, for different widths of the active zone :

$\Delta = 200$:

Table 4

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.3921	0.0287	0.8877
2	0.6079	0.7539	0.1123
3	-	0.2174	0

$\Delta = 2$:

Table 5

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.2177	0	0.8586
2	0.7823	0.5279	0.1414
3	-	0.4721	0

$\Delta = 2000$:

Table 6

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.4874	0.0329	0.8803
2	0.5126	0.7453	0.1197
3	-	0.2218	0

Comparing in respective order tables 4,5,6 with Tables 1,2,3 (where $\upsilon = 12$) it is clear that the probability of lower prices in the case of $\upsilon = 7$ have been significantly increased for all values of Δ. This agrees with the intuition that as the money flow into the market decreases, the probability of lower prices increases. This conclusion holds for all sellers psychological types.

5.2. "Reference-Price" tactics.

Let the reference price for all buyers be the same :

$$h^j = 130 \quad , \quad j=1,..,12$$

and consider the other market parameters unchanged. For Δ = 200, and γ_0 = 0.1 the following simulation results have been obtained.

5.2.1. Objective sellers.

With objective type of sellers we obtained :

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.0689	0.0010	0.6722
2	0.9311	0.5097	0.3278
3	-	0.4893	0

For this case we have also computed the absolute mean values of the utility function versus time in order to demonstrate the learning capability of the sellers automata. A plot of the optimality criterion for the second seller $\bar{Q}^2$ is shown in Fig. 4 (each point is the mean of 100 trials).

5.2.2. Hazardous seller's.

With hazardous type of sellers we obtained :

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.0760	0	0.3292
2	0.9240	0.1903	0.6708
3	-	0.8097	0

5.2.3. Cautious sellers.

With cautious type of sellers we obtained :

$E\{p_k^i[100]\}$ k \ i	1	2	3
1	0.3828	0.0006	0.9338
2	0.6172	0.7947	0.0662
3	-	0.2047	0

Comparing these results with the results of sec. 5.1 we see that they differ slightly.

Changing the reference prices h^j did not bring significant changes to the equilibrium probabilities.

5. CONCLUSIONS.

A model of many goal-oriented stochastic automata is introduced for the study of a certain class of problems of operations research. That class is characterized by the existence of a definite mutual opposition in the behavior of the participants in the collective. In the model the goals of the participants are assumed to be known only up to certain indeterminate parameters for which there is no a priori information available. Such class of problems cannot be solved by the theory of N-person games. By means of that automata model a numerical solution to the behavioral dynamics and the equilibrium conditions of the participants in the collective can be obtained.

Besides the automata model can demonstrate the effect of certain interesting factors like participants psychology, stimulation laws, behavioral tactics, etc... on the mode of collective behavior. As example, the model is used for the simulation of the price formation process in a free market. The result obtained demonstrate the applicability of the present automata model.

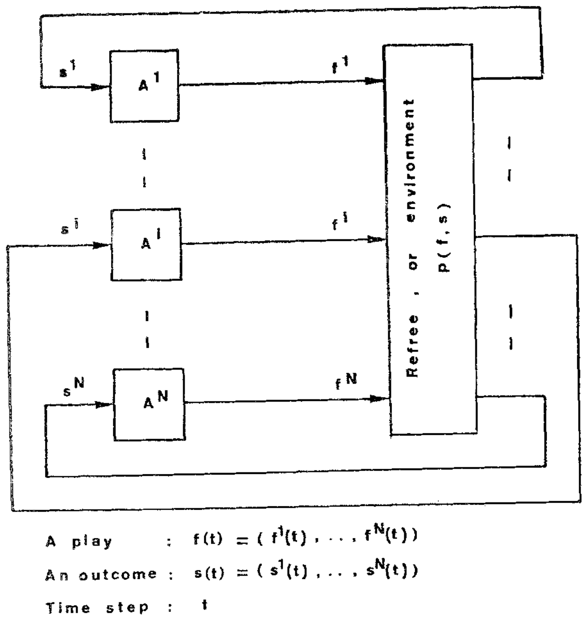

Fig. 1. N-automaton game.

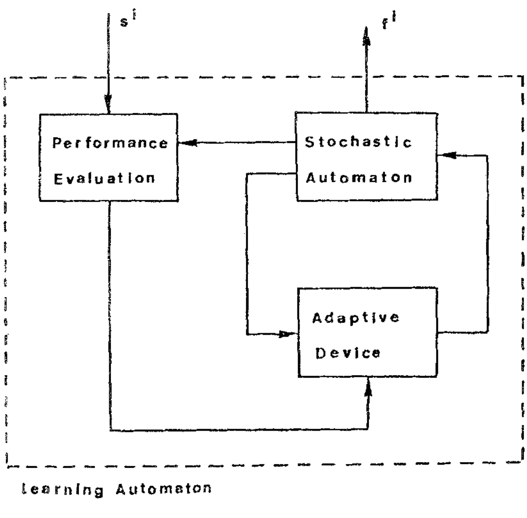

Fig. 2. Learning automaton.

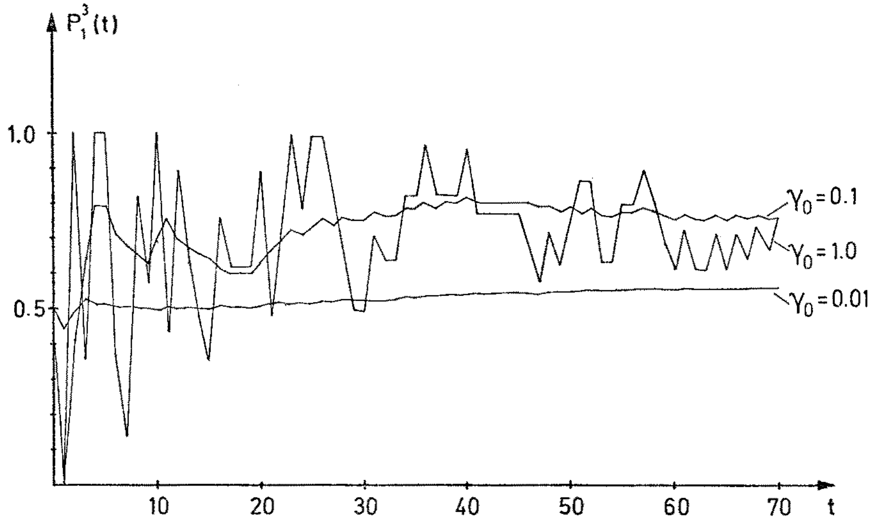

Fig. 3. Third seller first price probability versus time.

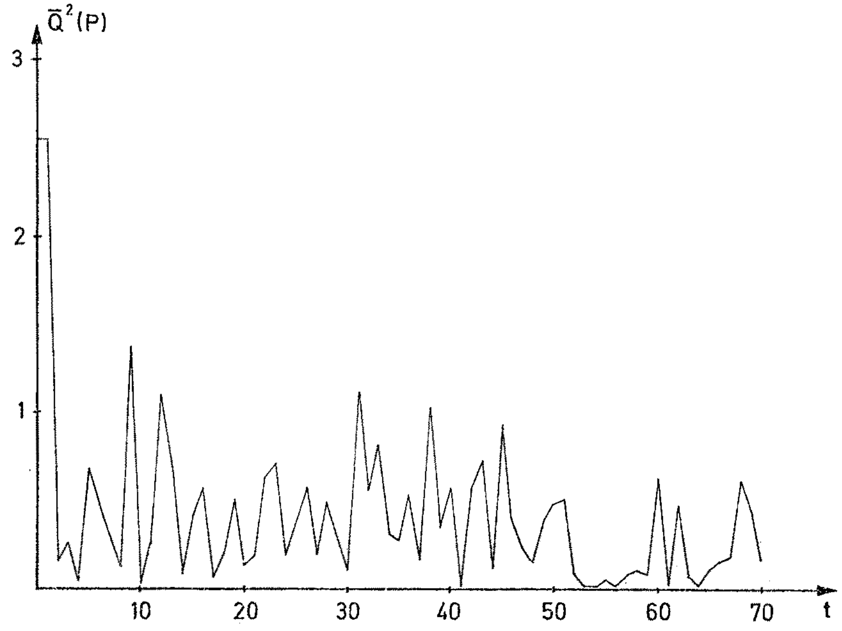

Fig. 4. Average magnitude of mismatch between supply and demand of the second seller commodity.

REFERENCES

1 L. Tsetlin, "On the Behavior of Finite Automata in Random Media", Automation and Remote Control, 22, p. 1210-1219 (1961).

2 V. Yu.Krylov, and M.L.Tsetlin., "Game between Automata", Automation and Remote Control, 24, pp. 889-899 (1962).

3 V.I.Varshavskii, and I.P.Vorontsova., "On the Behavior of Stochastic Automata with Variable Structure , Automation and Remote Control, 24, pp. 327-333 (1963).

4 K.S.Narendra, and M.L.A.Thathachar., "Learning Automata - A Survey", IEEE Trans., SMC-4, pp. 323-334 (1974).

5 Yu.D.Tenisberg., "Some Models of Collective Behavior in Dynamic Processes of Market Price Formation", Automation and Remote Control, n° 7, pp. 1140-1148 (1969).

6 A.V.Malishevskii, and Yu.D.Tenisberg., "One Class of Games Connected with Models of Collective Behavior," Automation and Remote Control, n° 11, pp. 1828-1837 (1969).

7 A.V.Malisheskii., "Models of Joint Operation of Many Goal-Oriented Elements,I,II, " Automation and Remote Control, n° 11,12, pp.1828-1845, 2020-2028 (1971).

8 S.Karlin., "Mathematical Methods in Games Theory, Programming, and Economics, I, Reading, Massachussetts, Addison-Wesley (1959).

9 V.I.Varshavskii, M.V.Meleshina, and V.T.Perekrest., "Use of Model of Collective Behavior in the Problem of Resources Allocation", Automation and Remote Control, n° 7, pp. 1107-1114 (1968).

10 L.P.Krylatykh., "On a Model of Collective Behavior",Engineering Cybernetics, pp. 803-808 (1972).

* * * *
* * *
* *