

ON ONE CONSTRUCTION OF A SEQUENCE OF AUTOMATA AND ITS BEHAVIOR IN GAMES

1964

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196401.99181>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

CYBERNETICS AND CONTROL THEORY

V. I. KRINSKII

ON ONE CONSTRUCTION OF A SEQUENCE OF AUTOMATA AND ITS BEHAVIOR IN GAMES

(Presented by Academician M. V. Keldysh, 25 I 1964)

In this note a construction is described of an asymptotically optimal sequence of automata, and games of two such automata with zero sum are considered. It is shown that if the memory capacities of both playing automata grow without bound, then the limit of the payoff lies between the upper and lower values of the game. If, moreover, there exists a limit of the ratio of the memory capacity of the second automaton to the memory capacity of the first, then the limiting payoff of the first automaton is a monotonically increasing function of the value of this ratio.

Problems on the behavior of automata in random environments and games are described in the work of M. L. Tsetlin ⁽¹⁾. We reproduce here the basic definitions from that work.

1. Let the input variable of the finite automaton $A_{M,n}$ be able to take only two values: $s(t) = 1$ (win) and $s(t) = -1$ (loss). The automaton $A_{M,n}$ has Mn states, and its output variable $f(t)$ can take one of M values $1, \dots, M$. The value $f(t)$ is called the action or strategy of the automaton at the moment t .

We shall say that the automaton $A_{M,n}$ functions in a stationary random environment $C(p_1, \dots, p_M)$ if the action i , produced at the moment t , generates at the moment $t + 1$ a loss with probability p_i and a win with probability $q_i = 1 - p_i$.

A sequence of automata $A_{M,n_1}, \dots, A_{M,n_k}, \dots$ is called asymptotically optimal in the environment $C(p_1, \dots, p_M)$ if, as $n \rightarrow \infty$, the mathematical expectation of the automaton's payoff in this environment tends to the maximum possible value

$$M(A_{M,n}; C) \rightarrow \max_i (q_i - p_i).$$

2. We shall describe the construction of a sequence of automata $D_{M,n}$, asymptotically optimal in an arbitrary stationary random environment ⁽²⁾. The automaton $D_{M,n}$ has M actions and Mn states. In the states $\varphi_1^k, \dots, \varphi_n^k$ this automaton performs action k ($k = 1, \dots, M$). The transitions between

the states of the automaton $D_{M,n}$ (for the case $M = 2$) are shown in Fig. 1.

Fig. 1

Upon a win, the states φ_α^i ($\alpha = 1, \dots, n$) pass into the states φ_n^i . Upon a loss, the states φ_α^i ($\alpha \neq 1$) pass into the states $\varphi_{\alpha-1}^i$, while the states φ_1^i , upon a loss, pass with equal probabilities M^{-1} into the states φ_n^k ($k = 1, \dots, M$).

It can be shown that the mean time during which the automaton $D_{M,n}$ performs action i is

$$T(p_i) = \frac{1 - p_i^n}{p_i^n}. \quad (1)$$

From this formula the asymptotic optimality of $D_{M,n}$ follows at once.

3. In a game of two automata with zero sum (¹), the outcome of each play (i, j) is a win for the first automaton (a loss for the second) with probability q_{ij} , and a loss for the first automaton (a win for the second) with probability $p_{ij} = 1 - q_{ij}$. Thus, in each play one of the automata wins and the other loses, and the game may be specified by the matrix $A = \|a_{ij}\| = \|q_{ij} - p_{ij}\|$ of the mathematical expectations of the payoff of the first automaton.
4. Consider the zero-sum game Γ of two automata D_{M,n_1} and D_{M,n_2} . This game is ergodic if the game matrix has no rows consisting entirely of ones and no columns consisting entirely of minus ones. In what follows we shall assume that this condition is satisfied.

Let R_{ij} be the probability that the automata play the play (i, j) . Then the mathematical expectation of the payoff of the first automaton is

$$W(n_1, n_2) = \sum_{i,j} R_{ij} a_{ij}. \quad (2)$$

It can be shown that the upper limit of $W(n_1, n_2)$ as $n_1, n_2 \rightarrow \infty$ does not exceed the upper value of the game, while the lower limit of $W(n_1, n_2)$ is not less than the lower value of the game. If there exists a limit l of the ratio n_2 to n_1 , then there also exists the limit $W(l)$ of the quantity $W(n_1, n_2)$, which is called the limiting payoff of the first automaton.

Let q_0 be the positive root of the equation $q_0^l + q_0 - 1 = 0$. Denote*

$$d = d(l) = 2q_0 - 1. \quad (3)$$

The following assertions hold.

- 1°. If the game Γ is such that $\max_i \min_j a_{ij} \geq d$, then $W = \max_i \min_j a_{ij}$.

2°. If $\min_j \max_i a_{ij} \leq d$, then $W = \min_j \max_i a_{ij}$.

3°. In the case

$$\min_j \max_i a_{ij} > d > \max_i \min_j a_{ij},$$

the limiting payoff is close to d in the following sense.

By a renumbering of the rows and columns, put the matrix A in the form

$$A = \begin{pmatrix} C & P \\ N & B \end{pmatrix},$$

where B is a maximal submatrix of the matrix A such that all elements of P are greater than d , and all elements of N are less than d . The set of elements of the matrix A not belonging to the matrix B will be denoted by A_l .

* From (1) and (3) it follows that the quantity d has the following property: for $a_{ij} > d$,

$$\lim_{n \rightarrow \infty} \frac{T_2(q_{ij})}{T_1(p_{ij})} = 0,$$

where $T_1(p)$ is the mean time during which the first automaton ($T_2(p)$ the second) in a stationary random environment performs an action for which it is penalized with probability p . For $a_{ij} < d$ the ratio of these times grows without bound.

** It can be shown that if $a_{ij} \in \bar{A}_l$, then as $n \rightarrow \infty$ the total probability of transition to the play (i, j) tends to zero, while if $a_{ij} \in A_l$, then this probability remains finite.

Let $a_{i_1 j_1}$ be the least of the elements A_l not smaller than d , and let $a_{i_2 j_2}$ be the greatest of the elements A_l not exceeding d . Then *

$$a_{i_1 j_1} \geq W \geq a_{i_2 j_2}.$$

Let us make the following remark. Case 1° occurs when the matrix $A = \|a_{ij}\|$ has an entire row consisting of elements not smaller than d . In all plays where the first automaton uses the strategy corresponding to this row, it is more inertial than the second automaton. In this case the behavior of the first automaton coincides with the behavior prescribed by the cautious tactic of game theory: the first automaton chooses a strategy that gives it the greatest guaranteed payoff. The second automaton minimizes the payoff of the first automaton.

In case 2°, the second automaton turns out to be in the same conditions as the first automaton in case 1°.

In case 3°, the automata carry out “mutual pursuit.” Only those plays which are played over the maximal time contribute a nonzero amount to the limiting payoff.

Let us also note that the limiting payoff W is not a monotone function of the elements of the game matrix (in contrast to the price of a game in von Neumann’s sense). There are even cases in which, when all elements of the matrix $\|a_{ij}\|$ are increased, the magnitude of the payoff decreases. Examples are provided by the games Γ_1 and Γ_2 of two identical automata

$$\|a_{ij}\|^{(1)} = \begin{pmatrix} 0.2 & -0.5 \\ -0.4 & 0.3 \end{pmatrix}, \quad \|a_{ij}\|^{(2)} = \begin{pmatrix} 0.5 & -0.2 \\ -0.1 & 0.6 \end{pmatrix}.$$

For identical automata $l = 1$ and $d = 0$. In both games all elements a_{ij} belong to A_l , and the limiting payoff of the first automaton in the game Γ_1 is 0.2, while in the game Γ_2 it is -0.1 , although $a_{ij}^{(2)} > a_{ij}^{(1)}$.

Consider two examples of games of automata with different memory sizes.

Example 1. The limiting payoff of the automaton $D_{2,n}$ in the game Γ_1 with the automaton $D_{2,ln}$, for $l > \lg p_{21} / \lg q_{11} \approx 0.7$, is equal to $\min \max a_{ij} = 0.2$, and for $l < \lg p_{21} / \lg q_{11}$ is equal to $\max \min a_{ij} = -0.4^*$.

Example 2. Consider the game consisting in the automaton $D_{2N,n}$ “guessing” the action of its opponent $D_{2N,ln}$, in which $a_{ij} = 1 - 2\alpha|i - j| - 2\beta$; $\alpha, \beta > 0$, $(2N - 1)\alpha + \beta < 1$. Denote

$$b_k = \lg(\beta + \alpha k) / \lg(1 - \beta - \alpha k + \alpha) \quad (k = 1, \dots, N).$$

For $l < b_N$, the limiting payoff is

$$W = \max \min a_{ij} = 1 - 2\beta - 2\alpha N.$$

For $l > b_1$, the limiting payoff is

$$W = \min \max a_{ij} = 1 - 2\beta.$$

For $b_k > l > b_{k+1}$, the limiting payoff is

$$W = 1 - 2\beta - 2\alpha k \quad (k = 1, \dots, N - 1).$$

5. The proof of properties 1°-3° of the limiting payoff in the game of automata $D_{M,n}$ and $D_{N,ln}$ is carried out, as in (*), by computing the quantities R_{ij} —the probabilities that the automata play the play (i, j) .

* A more precise formulation is as follows. Denote

$$\delta(a, d) = \begin{cases} \left(\frac{1+a}{2}\right)^l - \frac{1-d}{2}, & \text{if } a \geq d, \\ \frac{1-a}{2} - \frac{1-d}{2}, & \text{if } a \leq d; \end{cases}$$

$$\delta_0 = \min_{a_{ij} \in A_l} \delta(a_{ij}, d).$$

Then

$$\delta(W, d) \leq \delta_0.$$

If the game Γ is such that all elements a_{ij} belong to A_l , then as $n \rightarrow \infty$ the only probabilities of playing that remain nonzero are those of plays (i_0, j_0) for which $\delta(a_{i_0 j_0}, d) = \delta_0$.

Denote by $r_{\alpha, \beta}^{ij}$ the probability that the first automaton is in state φ_{α}^i , and the second in state φ_{β}^j . It is easy to see that, in the Markov chain describing the behavior of the system, only those of the states $\varphi_{\alpha, \beta}^{ij}$ are essential for which either $\alpha = n$ or $\beta = ln$,

$$r_{\alpha\beta}^{ij} = 0 \quad (\alpha \neq n, \beta \neq ln). \quad (4)$$

Denote $r_{n, ln}^{ij} = r_{ij}$. From the description of the automaton $D_{M, n}$ and the definition of the game,

$$r_{1, ln}^{ij} = r_{ij} p_{ij}^{n-1}, \quad r_{n, 1}^{ij} = r_{ij} q_{ij}^{ln-1}. \quad (5)$$

Consider the group of states corresponding to the play (i, j) . The probability that the system leaves this group of states is equal to the probability that the system enters it. The probability that the system leaves this group of states is

$$r_{1, ln}^{ij} p_{ij} + r_{n, 1}^{ij} q_{ij} = r_{ij} (p_{ij}^n + q_{ij}^{ln}).$$

The probability that the system enters this group of states is equal to

$$\frac{1}{M} \sum_u r_{uj} p_{uj}^n + \frac{1}{N} \sum_v r_{iv} q_{iv}^{ln}.$$

Therefore

$$r_{ij} = \frac{\tau_i + \sigma_j}{p_{ij}^n + q_{ij}^{ln}}, \quad (6)$$

where $\tau_i = \frac{1}{N} \sum_v r_{iv} q_{iv}^{ln}$, $\sigma_j = \frac{1}{M} \sum_u r_{uj} p_{uj}^n$.

Multiplying (6) by q_{ij}^{ln} and summing over j , we obtain (7), and multiplying by p_{ij}^n and summing over i , we obtain

$$\tau_i \sum_j \mu_{ij} = \sum_j \sigma_j \nu_{ij}, \quad (7)$$

$$\sigma_j \sum_i \nu_{ij} = \sum_i \tau_i \mu_{ij}, \quad (8)$$

where

$$\mu_{ij} = \frac{p_{ij}^n}{p_{ij}^n + q_{ij}^{ln}}, \quad \nu_{ij} = \frac{q_{ij}^{ln}}{p_{ij}^n + q_{ij}^{ln}}.$$

Studying the system (7), (8) as $n \rightarrow \infty$, from equation (6) one can obtain assertions 1°–3°.

6. In this note we have considered zero-sum games of automata of construction D . Apparently, assertions 1°–3° are also valid for a broader class of automata. For them the quantity d , determining the limiting payoff, should be computed from the equations

$$\lim_{n \rightarrow \infty} \frac{T_2(q_0)}{T_1(p_0)} = 1, \quad d = q_0 - p_0.$$

Institute of Biological Physics
Academy of Sciences of the USSR

Received
18 I 1964

CITED LITERATURE

1. M. L. Tsetlin, UMN, 18, No. 4 (1963).
2. V. I. Krinskii, Biophysics, 9, 4 (1964).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.