The principles of adaptive control of a technical system using Cetlin probabilistic automata

V I Petrenko, F B Tebueva, V O Antonov, V N Smykova and V E Nechvoloda
North-Caucasus Federal University, Stavropol, Russia

E-Mail: ant.vl.02@gmail.com

Abstract. This article examined the principles of adaptive control of a technical system using Cetlin probabilistic automaton. The study aims to reduce the learning time of the machine through the introduction of a progressive method for predicting short time series by the refined Brown method.

1. Introduction
Modern technical systems are faced with the tasks of intensive development and improvement of management methods. Today, artificial intelligence is used in many gadgets that make everyday life much easier and more convenient. In addition, artificial intelligence is increasingly used to solve complex social problems. These tasks include the development and implementation of adaptive control systems.

The effectiveness of the task assigned to the technical system depends on its state. It is proposed to consider the transition of the internal states of the system according to a mechanism with a given sequence of operations — an automaton. If the automaton is stochastic, then it does not act according to a predetermined rule, but adapts to environmental changes, changing the parameters of the system. One of the simplest and most effective forms of adaptive behavior of technical systems is the conditioned-reflex model of actions based on the training of probabilistic automata. This is a mechanism for constantly changing the state of an automaton according to the results of past experience over time to perform a formalized task.

In the presented works [1,2], the learning process of a probabilistic automaton is considered, which allows to make the optimal decision through repeated interaction with the environment. However, the learning process can be simplified using the method presented in this paper. There is another result of the learning theory of a deterministic probabilistic automaton [3]. However, the method used in this source is impractical to apply to this article.

The article [4] describes the prediction of the future process of the system depending on previous tasks or requests using controlled learning. The use of forecast values for calculating the efficiency of the power system is given in the source [5]. The proposed solution can be used to modify the results for adaptive energy efficiency management of a technical system.

In this paper [6], we propose an improvement in the learning process of the conditioned-reflex behavior model of a technical system based on Cetlin probabilistic automata using progressive forecasting techniques to reduce the number of systems learning steps using the refined Brown method.
2. Study of a technical system based on probabilistic Tsetlin automata

2.1. Mathematical statement of the problem
There is an automaton, inside its states the transition occurs randomly. [7]:

\[ A = (X, Y, S, P, \mathcal{E}^0) \]

where \( X, Y \) - input and output signals, respectively;
\( S \) - state matrix of the automaton, where the set of states has the form \((s_1, \ldots, s_n)\);
\( P \) – automaton behavior, random function of the form \( P: S \times X \rightarrow S \times Y \), moreover \( \forall (s, x, s', y) \in S \times X \times S \times Y \Rightarrow P = (s, x, s', y) \).
\( \mathcal{E}^0 \) – the initial distribution of the probabilistic automaton, the probability that the BA is in a state \( s \), moreover \( \forall s \in S \), i.e. the probability that the machine is in a state \( S \) at the initial time \((t = 0)\).
The actions of the automaton occur with a stochastic three-dimensional matrix \( S \) of size \( S_k \times X \times Y \), where \( S_k \) – is the number of states of the object, \( X \) \& \( Y \) – are the input and output signals.

2.2. Disclosure of the method for solving the problem
When receiving the input action Score automaton reaction occurs. The system is punished or promoted, which affects the penalty function, which changes the likelihood of the transition of the machine from state to state.

Figure 1 shows a diagram of a simple automaton (an automaton \( A \)) with linear tactics having five states in two modes of behavior, where \( C > 0 \) – the first mode, and \( C < 0 \) – the second mode, with \( C \) – the state number of the automaton.

![Figure 1. Scheme of a simple automaton with linear tactics with 5 states in 2 modes of behavior.](image)

Arrows with shaded tips are actions / changes of the machine when it is promoted, and arrows with empty ones - when it is punished. Arrows with shaded tips are actions / changes of the machine when it is promoted, and arrows with empty ones - when punished. The probabilities of transition from one condition to another are denoted as \( p_{c_a \rightarrow c_b} \), where \( c_a \) \& \( c_b \) – are the adjacent conditions of the automaton. The automaton is supplied with a matrix of probabilities of transitions of the system by state (table 1).

| Condition | -2 | -1 | 0 | 1 | 2 | ... | b |
|-----------|----|----|---|---|---|-----|---|
| -2        | \( p_{c_{-2},c_{-2}} \) | \( p_{c_{-2},c_{-1}} \) | 0 | 0 | 0 | ... | \( p_{c_{-2},c_b} \) |
| -1        | \( p_{c_{-2},c_{-1}} \) | 0 | \( p_{c_{-1},c_0} \) | 0 | 0 | ... | \( p_{c_{-1},c_b} \) |
| 0         | 0  | \( p_{c_0,c_{-1}} \) | \( p_{c_0,c_{-2}} \) | \( p_{c_0,c_1} \) | 0 | ... | \( p_{c_0,c_b} \) |
| 1         | 0  | 0  | \( p_{c_1,c_{-2}} \) | 0 | \( p_{c_1,c_2} \) | ... | \( p_{c_1,c_b} \) |
| 2         | 0  | 0  | 0  | \( p_{c_2,c_{-1}} \) | \( p_{c_2,c_1} \) | ... | \( p_{c_2,c_b} \) |
| ...       | ...| ...| ...| ...| ...| ... | ... |
| a         | \( p_{c_a,c_{-2}} \) | \( p_{c_a,c_{-1}} \) | \( p_{c_a,c_0} \) | \( p_{c_a,c_1} \) | \( p_{c_a,c_2} \) | ... | \( p_{c_a,c_b} \) |

Table 1. Transition probability matrix for conditions of the automaton \( A \).
If at the initial moment of time the machine receives a promotion from a position \( C = 0 \), then the system seeks to maintain this mode with probability \( P_{c_0,c_0} \). When being punished, it passes with probability \( P_{c_0,c_1} \) or \( P_{c_0,c_{-1}} \) to one of the vertices \( C = -1 \) or \( C = 1 \), from which, upon receipt of the reaction of the system, it passes to another adjacent state, etc. The implementation of inertia is based on a change in the probability \( p_{c_a,c_b} \) of occurrence of events.

2.3. Advanced training algorithm for the Tsetlin linear automaton by the refined Brown method

Let’s represent the binary time series of the reaction of the system as a sequence of numbers representing the number of identical reactions of the system to the input action, i.e. \( W_{шифр} = \{w_1^n, w_2^n, ..., w_k^n\} \), where \( w_k^n = \sum_i w_i \), where \( w_i = w_{i+1} \). For example: \( W_{шифр} = 142323 \ldots \), where the first value of “1” indicates the number of promotions of the machine in a row, and the number “4” indicates the number of punishments. Next in order. To draw up a predictive model, it is necessary to ascertain that the chain of identical system responses to the input action is terminated, i.e. \( w_i \neq w_{i+1} \).

We study this series for the presence of a persistence property, i.e. the value of the Hurst indicator \( \tilde{H} \) should be in the range \([0;7;1)\). A persistent time series has a long-term memory; therefore, it has long-term correlations between current events and future events [8].

If the time series meets the proposed requirements, then based on the calculation of the averaged Hurst indicator \( \tilde{H} \) and the use of the refined Brown method, we continue the studied time series with the forecast value according to the formula:

\[
W_{k+1}^n = D \cdot W_k^n + (1 - D) \cdot W_{k-1}^n,
\]

(2)

Where \( D \) – fractal time series dimension, \( D = 2 - \tilde{H} \).

Given the average forecasting error of method \( \approx 5.4\% \) we have a further scenario for the development of the reaction of the automaton. The predicted value is nothing but the next value of the number of punishments or promotions of the system \( w_k^n \).

2.4. Development of a training method for an automaton with linear tactics through the introduction of a forecasting technique

As an example of the application of the proposed training method, we consider a simple automaton with linear tactics \( A_1 \), 4 possible conditions: \( C=[0;1a;1b] \) (figure 2) and a given probability matrix of transitions over the conditions of the automaton \( A_1 \) (table 2).

For training the machine, a penalty function has been introduced depending on the conditions of the experiment; for the state \( C = 0 \) fine \( S_{h_0} = 0.01 \), For the condition \( C = 1 \) fine \( S_{h_1} = 0.02 \), for the conditions \( C = 1a; 1b: S_{h_{1a};1b} = 0.005 \). I.e. when punishing an assault rifle in a condition \( C \) probabilities units \( S_{h_c} \) are fined in a state in the selected direction, and added to an alternative favorable direction. With encouragement, the procedure is repeated exactly the opposite. The penalty is entered at the end of the cycle and the object returns to the “0” state.

Figure 2. A1 assault rifle with linear tactics 4 states.
Encouraging directions of instruction in this machine are highlighted by a filled arrow, while punishment corresponds to empty arrows.

Table 2. The matrix of transition probabilities for the conditions of the automaton.

| Condition | 0   | 1   | 1a  | 1b  |
|-----------|-----|-----|-----|-----|
| 0         | 0.5 | 0.5 | 0   | 0   |
| 1         | 0   | 0   | 0.5 | 0.5 |
| 1a        | 0.5 | 0   | 0.5 | 0   |
| 1b        | 0.5 | 0   | 0   | 0.5 |

The initial position of the system is in the state “0”. After 129 learning steps, a short time series of 40 values was obtained, 20 punishments and rewards under even and odd numbers, respectively (table 3).

Table 3. Time series of the number of identical reactions of the automaton to input actions.

| №  | Wч,1 |
|----|-------|
| 1  | 1     |
| 2  | 3     |
| 3  | 4     |
| 4  | 6     |
| 5  | 7     |
| 6  | 8     |
| 7  | 9     |
| 8  | 10    |
| 9  | 11    |
| 10 | 12    |
| 11 | 13    |
| 12 | 14    |
| 13 | 15    |
| 14 | 16    |
| 15 | 17    |
| 16 | 18    |
| 17 | 19    |
| 18 | 20    |
| 19 | 21    |
| 20 | 22    |
| 21 | 23    |
| 22 | 24    |
| 23 | 25    |
| 24 | 26    |
| 25 | 27    |
| 26 | 28    |
| 27 | 29    |
| 28 | 30    |
| 29 | 31    |
| 30 | 32    |
| 31 | 33    |
| 32 | 34    |
| 33 | 35    |
| 34 | 36    |
| 35 | 37    |
| 36 | 38    |
| 37 | 39    |
| 38 | 40    |

Let us consider the tendency for the number of promotions to change on the graph (figure 3), where on the abscissa axis the number of the group of identical reactions is displayed, and on the ordinate axis the value of the number of promotions.

Figure 3. Schedule of the number of machine promotions \( A_1 \).

Analyzing the time series in figure 3, it should be noted that the number of rewards has increased markedly over time, which proves the adaptability of the machine. Consider the new matrix of probabilities of the automaton after 129 steps (table 4).

Table 4. Probability matrix of the automaton \( A_1 \) after 129 training steps.

| Condition | 0   | 1   | 1a  | 1b  |
|-----------|-----|-----|-----|-----|
| 0         | 0.22| 0.78| 0   | 0   |
| 1         | 0   | 0   | 0.1 | 0.9 |
| 1a        | 0.455| 0   | 0.545| 0   |
| 1b        | 0.14| 0   | 0   | 0.86 |

Based on the data given in the table, it should be concluded that the previous choice of the direction of movement of the system from two equally probable values was changed in such a way that the probability of choosing a more favorable state (1b) increased, thereby the machine adapts to the established conditions.

We calculate the predicted values of the time series using the updated Brown method. We predict the next value of the time series using the averaged Hurst indicator and the refined Brown model. The result is displayed on the table. 5.
Table 5. Comparison of the 41st real and forecast value of the number of system promotions.

| №41 | Real | Predictive | Difference | Unit/ % |
|-----|------|------------|------------|---------|
| W_{ch}^{41} | 10 | 9 | 1/ 10% |

Based on the data table 5 it should be concluded that the error between the real and the predicted value is 10%. This value corresponds to the characteristic of the probability of a condition transition С = 1b, therefore it can be concluded that the next cycle of this measurement error will be extinguished by changing the probabilities of the condition of the system at each input to the condition С = 0.

3. Results

Using the predicted value in the learning model of the machine, we compare the percentage of the benefits of the time saved after applying the first insert of the predicted value according to the formula

\[ \Delta \tilde{\delta}_{\text{reduction}} = 100\% \ast \left(1 - \frac{\tilde{\delta}_{\text{forecast}}}{\tilde{\delta}_{\text{classic}}}\right) \]  \hspace{1cm} (3)

where \( \tilde{\delta}_{\text{classic}} \) – time spent on training the machine using the classical method until the end of the training \( W_{ch,real}^{41} \) step corresponding to the predicted one (4):

\[ \tilde{\delta}_{\text{classic}} = \sum_{n=1;41} W_{ch}^{n} \ast t; \]  \hspace{1cm} (4)

\( \tilde{\delta}_{\text{forecast}} \) – time spent on training the machine, taking into account the first predicted value

\[ \tilde{\delta}_{\text{forecast}} = \sum_{n=1;41} W_{ch}^{n-1} \ast t \]  \hspace{1cm} (5)

To compare the characteristics, we assume that the response of the system to one input action will be obtained through \( t = 1 \) sec. Then the potential time savings on the 41st cycle, corresponding to the 130-138 step of training, will be

\[ \Delta \tilde{\delta}_{\text{reduction}} = 100\% \ast \left(1 - \frac{129}{138}\right) = 6,52\%. \]  \hspace{1cm} (6)

Given the huge input and output alphabets \( X, Y \), the use of a predictive model on every second promotional element of the time series of the number of identical system reactions to the input effect will reduce the system learning time

\[ \Delta t_{\text{reduction}} = \tilde{\delta}_{\text{full}} - \tilde{\delta}_{\text{forecast}}^\text{max} \]  \hspace{1cm} (7)

where \( \tilde{\delta}_{\text{full}} \) – full time of learning the machine by the classical method (8):

\[ \tilde{\delta}_{\text{full}} = \sum_{n=1;\text{max}} W_{ch}^{n} \ast t; \]  \hspace{1cm} (8)

\( \tilde{\delta}_{\text{forecast}}^\text{max} \) – training time won by all inserts of predictive steps in the training steps (9):

\[ \tilde{\delta}_{\text{forecast}}^\text{max} = \left(\sum_{n=1;\text{max}} W_{ch}^{n} - \sum_{n=1;41} W_{ch,\text{forecast}}^{n}\right) \ast t \]  \hspace{1cm} (9)

Within the framework of this experiment, we assume that the complete learning process of the automaton includes the total \( \sum_{n=1;\text{max}} W_{ch}^{n} = 500 \) steps of the reaction values of the system (for comparison, only 41 steps are considered in table 2), then \( \tilde{\delta}_{\text{full}} \) for \( t = 1 \) sec, is equal to: \( \tilde{\delta}_{\text{full}} = t \ast \)
\( W_{\text{full}} = 500 \text{ sec}. \) Also, as part of the example, we assume that \( \sum_{n=1}^{41} W_{\text{ch,forecast}} = 150 \text{ steps}, \) then:

\[
\delta_{\text{max}}^{\text{forecast}} = (500 - 150) \times 1 = 350 \text{ sec}.
\]

In proportion to the growth in the total number of steps, the number of predictions increases. Thus, we conclude that the reduction in the learning time of the Cetlin probabilistic automaton changes by a value calculated by formula 4.

Consider the difference in the learning speed of a machine with the classic and advanced method:

\[
\Delta t_{\text{reduction}} = 500 - 350 = 150 \text{ sec}.
\]

So, under given conditions, the effectiveness of the improved method can be estimated as 30% time savings.

4. Conclusion

This paper presents the use of a predictive model for training technical systems based on probabilistic automata. The method proposed in the article provides less time for a technical system with adaptive control, in contrast to approaches that are not used in predicted values. The prediction allows to predict the future behavior of the machine with a certain degree of certainty to accelerate the learning process. According to the results of the experiment, the proposed method saves 30% of the time spent on training. Thus, the obtained results are expediently applied to improve the learning speed of technical systems with adaptive control.

It should be noted that this method is applicable in the modeling of more complex forms of adaptive behavior, which focuses on the adaptive properties inherent in both model and living organisms. Also, the proposed mechanism for using the predictive model in learning a probabilistic automaton has less computing power, but it has a small calculation error equal to \( \approx 10\% \), which is smoothed out during the next learning step, while reducing learning time in proportion to the number of learning steps, which makes it possible to use probabilistic machines in real time.

References

[1] Ge H, Jiang W, Li S, Li J, Wang Y and Jing Y 2015 A novel estimator based learning automata algorithm Kluwer Academic Publishers DOI: 10.1007/s10489-014-0594-1 2015

[2] Rezvanian A A, Saghihi A M, Vahidipour S M, Esnaashari M and Meybodi M R 2018 Learning automata theory Studies in Computationai Intelligence DOI: 10.1007/978-3-319-72428-7_1

[3] Balle B, Castro J and Gavaldà R 2013 Learning probabilistic automata: A study in state distinguishability Journal Theoretical Computer Science 46-60

[4] Gritsenko A V, Demurchev N G, Kopytov V V and Shulgin A O 2015 Decomposition analysis and machine learning in a workflow-forecast approach to the task scheduling problem for high-loaded distributed system DOI: 10.5539/maas.v9i38.93909

[5] Guo Z, Zhou K, Zhang X and Yang S 2018 A deep learning model for short-term power load and probability density forecasting DOI: 10.1016/j.energy.2018.07.090

[6] Tebueva F B 2016 An improved Brown method applying fractal dimension to forecast the load in a computing cluster for short time series Indian Journal of Science and Technology 9(19) DOI: 10.17485/ijts/2016/v9i19/93909

[7] Andrew M 2015 Mironov A theory of probabilistic automata, part 1 Cornell University Library: Cite 1507.05164

[8] Dobrynin D A, Karpov V E, Meshcheryakova T V and Stepanov S N 2005 Adaptive Mobile Robot Creative Scientific and Technical Laboratory of the Polytechnic Museum (Moscow) Access Mode: rema44.ru/resurs/papers/karpov/stat2005.doc