

тизированная система мониторинга позволяет нейтрализовать это воздействие. На массиве данных, накапливаемой системой, становится целесообразным проведение OLAP-анализа, дающего наглядную оперативную информацию в рамках педагогического мониторинга как в процессе очного обучения, так и при обучении дистанционными методами.

Библиографический список

1. Поташиник, М. М. Управление качеством образования / М. М. Поташиник. М. : Пед. о-во России, 2006.
2. Бондарева, О. В. Мониторинг качества обучения студентов как средство управления образованием / О. В. Бондарева // Вопросы интернет-образования [Электронный ресурс]. Электрон. журн. 2008. Вып. 56. Режим

доступа: http://vio.fio.ru/vio_56/cd_site/Articles/art_1_6.htm. Загл. с экрана.

3. Матрос, Д. Ш. Управление качеством образования на основе новых информационных технологий и образовательного мониторинга / Д. Ш. Матрос, Д. М. Полев, Н. Н. Мельникова. М. : Пед. о-во России, 2001.
4. Пантелеев, Е. Р. Оперативный анализ действий студента в среде web-обучения / Е. Р. Пантелеев, В. А. Суворов // Информ. технологии. 2008. № 4. С. 67–72.
5. Федоров, А. Введение в OLAP-технологии Microsoft / А. Федоров, Н. Елманова. М. : Диалог-МИФИ, 2002.
6. Товбис, Е. М. «Электронная лаборатория» – система поддержки учебного процесса / Е. М. Товбис, И. Я. Дорошенко // Проблемы информатизации региона : материалы десятой Всерос. науч.-практ. конф. / Сиб. федер. ун-т. Красноярск, 2007. Т. 2. С. 157–163.

Е. М. Tovbis

OLAP TECHNOLOGY IN EDUCATIONAL PROCESS MONITORING

The detailed estimation of student's progress is impossible without educational process monitoring at University. The technology of multidimensional data analysis in OLAP is offered to be used. The paper considers the example of using OLAP technology in educational process in disciplines limitation.

Keywords: monitoring, OLAP, educational process.

© Товбис Е. М., 2009

УДК 004.588

М. V. Karaseva

INFORMATION TRAINING TECHNOLOGY: LEARNER'S MEMORY STATE MODEL

This paper mainly studies the foreign language training problem concerned with the application of the learner's memory state model in the informational-training technology. One of the approaches to intellectual technologies synthesis in training, developing new methods of training, proposed by professor L. A. Rastrigin was used. An exponential dependency of ignorance probability of a lexical unit upon the forgetting rate and time was taken as a trainer model.

Keywords: IT-technology, learner's memory state model, training process, information portion, training algorithm.

The training process is one of the oldest ones in the world. Everybody played the role of learner and teacher, and everyone knows that it is rather difficult in both cases. For the problem of training foreign language we consider it is of greater importance. In contrast to other fields the process of training foreign language is based on the properties of human memory which is different amongst people.

The information computer systems are the main tool of modern training technique. This choice is absolutely clear: as a matter of fact the training process is a transfer of information, and information searching systems should do it. But it is one of possible approaches to training as information process. There exist fields where it is very effective, but this approach should not be considered

absolutely right, because it does not try to individualize the training process.

In recent years another approach is developed when training can be considered as control [1]. The relations between teacher and learner can be interpreted as relations between the control object and the control device. This approach allows to use the methods of control theory. For effective training it is necessary to have a trainer model. Surely it is an approximated model and we should adapt its parameters and structure to provide its adequacy to real object.

This approach is rather advanced for the problem of training foreign language because the trainer capability to master lexical units is connected with his (her) memory [2].

We modified the approach of professor L. A. Rastrigin in order to solve this problem in a multilingual sense. We will take a result of psychological memory research as a trainer model, namely exponential dependency of ignorance probability of a lexical unit upon forgetting rate and time. Further we are going to consider it in detail.

Training as control. The standard training problem usually consists in that a learner should memorize determined information portions, and this problem can be formulated as the control problem.

In this case the learner is an object of control, and a teacher or trained device is a source of control. Obviously, this is a complex object and we can use all known principles of control by a complex object [1].

Look at the training process scheme (fig. 1). The control object is an object of training (called a “trainer”), the control device is a trained unit (called a “teacher”). Y is a trainer’s state, measured by data unit D_y ; Y' – information about trainer’s state got by a teacher in answer to questions U , besides U includes the portions of training information. Purposes of training Z^* and resources R are given to the teacher.

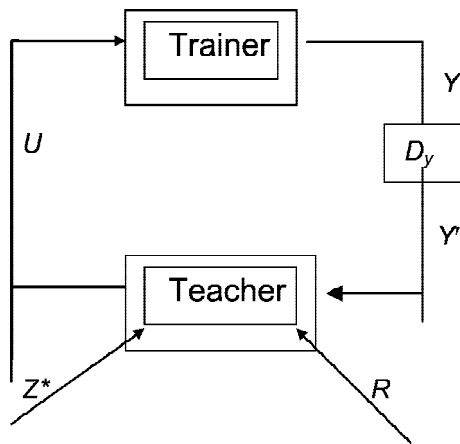


Fig. 1. Training process scheme

The problem is to organize the training U , changing trainer’s state Y to achieve the formulated purposes Z^* :

$$U = \varphi(Y, Z^*, R),$$

where φ is an algorithm of training.

As a rule the training process in computer training systems is realized in the next way. A trainer gets an information portion, which should be learnt. Then he is tested to determine a level of training quality. Answers of the trainer are checked and he gets the new information portion. The information portion sequence is determined by the training program [3]. The linear and branched training programs are widely spread. Besides the branched programs are divided in two types: outer and inner regulating ones.

New statement of the training problem. We should formulate new statements of the training problem [4]. We formalize:

- training purpose Z^* ;
- training information ;
- trainer model;
- algorithm of training .

We present the training process as sequence of seances (lessons) beginning at time moments t_0, t_1, \dots, t_n . To determine the effectiveness we take a function of training quality $Q_n = Q(Y_n)$, where Y_n – state of object at moment t_n . The purpose is to minimize Q by training:

$$Q(Y) \rightarrow \min_{u \in U}.$$

Since a level of absolute training Q^* can not be achieved because of real properties of the human memory the training process should be terminated, when the criterion of quality reaches some given quantity δ :

$$Q_n \leq \delta,$$

where $\delta > Q^*$ is quantity, close to Q^* .

Purpose Z^* can be formalized as

$$Z^*: \begin{cases} Q(Y^*) \leq \delta, \\ T(Y^*) = \min, \end{cases}$$

where $T(Y^*)$ – time, necessary to reach the state Y^* .

We formalize training information (TI). We consider such process, where TI can be presented as a finite state of enumerated elementary portions: $U = \{1, 2, \dots, N\}$. At every seance with help of the training algorithm the following subset is constructed:

$$U_n = \{u_1, u_2, \dots, u_{M_n}\}, u_i \neq u_j \text{ при } i \neq j, u_i \in U.$$

This subset contains M_n elementary portions of TI with the corresponding numbers.

We consider the trainer model. We describe a trainer state at the n -th seance by an ignorance probability vector:

$$Y_n = P_n = (p_1^n, p_2^n, \dots, p_N^n),$$

where p_i^n – ignorance probability of the i -th element at moment t_n .

Using data of psychology in the field of memory research we choose exponential dependency as a model:

$$p_i^n = p_i(t_i^n) = 1 - e^{-a_i^n t_i^n},$$

where a_i^n – rate of forgetting the i -th element at the n -th seance; t_i^n – time after last learning the i -th TI element. The forgetting rate of each element is reduced, if this element is given to the trainer to learn, and it is not changed otherwise:

$$\alpha_i^{n+1} = \begin{cases} \alpha_i^n, & \text{if } i \notin U_n, \\ \gamma' \alpha_i^n, & \text{if } i \in U_n \text{ and } r_i^n = 0, \\ \gamma'' \alpha_i^n, & \text{if } i \in U_n \text{ and } r_i^n = 1, \end{cases} n = 1, 2, \dots,$$

where $\gamma', \gamma'', \alpha_i^1$ ($i = 1, 2, \dots, N$) – parameters defining the individual features of trainer memory; $0 < \gamma' < \gamma'' < 1, \alpha_i^1 > 0$.

Answers to tests can be written as:

$$r_i^n = \begin{cases} 0, & \text{if answer is right,} \\ 1, & \text{otherwise.} \end{cases}$$

The effectiveness criterion Q should define a level of training. For the foreign language learning problem the level of training is defined by ignorance probability of any TI element:

$$Q_n = \sum_{i=1}^N p_i(t_i^n) q_i^{\max} \xrightarrow{U_n} \min \Rightarrow U_n^*,$$

where $0 < q_i < 1$ is a frequency of the element appearance in the given text, $Sq_i = 1$; $q_i^{\max} = \max q \{q_{i1}, q_{i2}, q_{i3}\}$, q_{i1}, q_{i2}, q_{i3} – frequencies of English, German and Russian word from the corresponding multilingual frequency dictionary [5].

A result of the problem solution is an optimal TI portion U_n^* given at the n -th seance.

To minimize Q_n it is necessary to include the elements of the greatest meaning of multiplication $p_i(t_i^n)q_i$ in U_n^* because their memorizing causes multiplication vanishing and decreases the meaning Q_n essentially.

Therefore it is required to find M_n maximal members of sum in the criterion, whose indexes define TI portion. They can be found by the rule:

$$\begin{aligned} u_1 &= \arg \max_{1 \leq i \leq N} p_i(t_i^n)q_i^{\max}, \\ u_2 &= \arg \max_{1 \leq i \leq N, i \neq u_1} p_i(t_i^n)q_i^{\max}, \\ &\dots \\ u_{M_n} &= \arg \max_{1 \leq i \leq N} p_i(t_i^n)q_i^{\max}, \\ i &\neq u_j \quad (j = 1, 2, \dots, M_n - 1), \end{aligned} \quad (1)$$

where $\arg \max \{a_i\} = i^*$ is index $i^* \in U$ of maximal a_i , it means $a_{i^*} = \max a_i$, and $\{u_1, \dots, u_{M_n}\} = U^* - \text{TI portion given to the trainer at the } n\text{-th seance.}$

Let T_n be the duration of the n -th seance or time for learning the portion. We assume that time of learning the i -th element is directly proportional to its ignorance probability. Then

$$M_n = \max_{1 \leq M \leq N} \left\{ M : T_n \geq \kappa \sum_{i \in \{u_1, \dots, u_M\}} p_i(t_i^n) \right\}, \quad (2)$$

where κ – average learning time of TI element at its first presentation to the trainer; u_1, \dots, u_{M_n} – numbers of TI elements. Parameter κ is unknown a priori and should be adapted:

$$\kappa_{n+1} = \kappa_n + v (T'_n - T_n), \quad (3)$$

where v is unmeasured coefficient of adaptation rate, T'_n is time spent by the trainer to learn U_n .

Training Algorithm. Thus the training algorithm [6] can be presented in the following way.

1. Check knowledge of portion U_n and construct the set R_n .

2. Realize parameter adaptation:

$$\alpha_i^{n+1} = \begin{cases} \alpha_i^n, & \text{if } i \notin U_n, \\ \gamma' \alpha_i^n, & \text{if } i \in U_n \text{ and } r_i^n = 0, \\ \gamma'' \alpha_i^n, & \text{if } i \in U_n \text{ and } r_i^n = 1, \quad n = 1, 2, \dots, \end{cases}$$

3. Correct the ignorance probability vector, i. e. form P_{n+1} according to our exponential dependency, using previous rule and taking into account forgetting time of information after last learning t_i^n :

$$t_i^{n+1} = \begin{cases} \Delta t_n, & \text{if } i \in U_n, \\ t_i^n + \Delta t_n, & \text{if } i \notin U_n, \quad n = 0, 1, \dots \end{cases}$$

4. Calculate Q_{n+1} .

5. If $Q_{n+1} \leq \delta$ then the training process is terminated; otherwise you should define portion U_{n+1} by (1), and present this portion to the learner taking into account rules (2) and (3). The points 1...5 are repeated again and so on.

Model Parameter Estimate. When exploring the trainer model the problem of estimate of unknown parameters is appeared. We should estimate parameters of correction of forgetting rates γ' and γ'' and initial meaning of these rates $\alpha^1 = (\alpha_1^1, \dots, \alpha_N^1)$. These parameters describe individual features of the trainer memory.

α^1 can be estimated by maximal likelihood method. We can show it [3]. The likelihood function is constructed:

$$P = p^x (1-p)^{K-x},$$

where $x = \sum_{i \in U_n} r_i^n$ is a number of elements unremembered from K_n . The set $U_n = \{u_1, \dots, u_{K_n}\}$ contains the elements given to the trainer for the first time. Changing P for $\ln P$, we have

$$\frac{\partial \ln P}{\partial \alpha} = \frac{x e^{-\alpha t}}{1 - e^{-\alpha t}} + (x - K_n)t = 0.$$

We get an estimate of parameter α :

$$\hat{\alpha} = -\frac{1}{t} \ln \frac{K_n - x}{K_n}.$$

Parameters γ' and γ'' can be also estimated by maximal likelihood method [3]. Before training we should have experimental data presented by the set R_n . The trainer gets N strange words to memorize. Then he is tested, the required set R_n contains the test results. The trainer should learn the unremembered words until $r_i^n = 0$ for all $i = 1, 2, \dots, N$.

The data obtained with help of this experiment can be described by our model. In this case $M_n = N$ and $\Delta t_n = 1$ for all $n = 0, 1, \dots, K$, where K – number of tests, i. e.

$$K = \min \left\{ n : \sum_{i=1}^N r_i^n = 0 \right\}.$$

The forgetting rates are changed

$$\alpha_i^{n+1} = \begin{cases} \gamma' \alpha_i^n, & \text{if } r_i^n = 0, \\ \gamma'' \alpha_i^n, & \text{if } r_i^n = 1, \end{cases}$$

where $0 < \gamma' < \gamma'' < 1$, $i = 1, 2, \dots, N$, $n = 1, 2, \dots, K$. The ignorance probabilities have the following form:

$$p_i^n = 1 - e^{-\alpha_i^n}, \quad i = 1, 2, \dots, N, \quad n = 1, 2, \dots, K.$$

It is taken into account here that all the words ($M_n = N$) are learned everyday ($\Delta t_n = 1$).

Let A_n be the set of word numbers unremembered before the n -th test; S_n – amount of these words, $|A_n| = S_n$. Then for all $i \in A_n$ we have

$$\alpha_i^n = (\gamma'')^{n-1} \alpha_i,$$

$$p_i^n = 1 - e^{-\alpha_i^n} = 1 - e^{-(\gamma'')^{n-1} \alpha_i} = p_i, \quad n = 1, 2, \dots$$

The experimental data can be presented by the set R_n of realizations r_i^n of variates ξ_i^n with the following distribution:

$$P\{\xi_i^n = 1\} = p_i^n,$$

$$P\{\xi_i^n = 0\} = 1 - p_i^n, \quad i = 1, 2, \dots, N.$$

We denote

$$x_n = \sum_{i \in A_n} (1 - r_i^n),$$

and construct the likelihood function

$$\begin{aligned} P &= \prod_{n=1}^K p_n^{S_n - x_n} (1 - p_n)^{x_n} = \\ &= \prod_{n=1}^K [1 - e^{-(\gamma'')^{n-1} \alpha}]^{S_n - x_n} [e^{-(\gamma'')^{n-1} \alpha}]^{x_n}. \end{aligned}$$

We change P for $\ln P$ and take a partial derivative over γ' in order to find a maximal meaning $\ln P$:

$$\frac{\partial \ln P}{\partial \gamma''} = \sum_{n=2}^K \left[\frac{S_n - x_n}{1 - e^{-(\gamma'')^{n-1} \alpha}} e^{-(\gamma'')^{n-1} \alpha} (n-1) \times \right. \\ \left. \times (\gamma'')^{n-2} \alpha - x_n (n-1) (\gamma'')^{n-2} \right] = 0.$$

We get an equation for finding γ'

$$\sum_{n=2}^K \frac{S_n - x_n}{1 - e^{-(\gamma'')^{n-1} \alpha}} e^{-(\gamma'')^{n-1} \alpha} (n-1) (\gamma'')^{n-2} = \\ = \sum_{n=2}^K x_n (n-1) (\gamma'')^{n-2}.$$

Obviously, it is impossible to get an exact solution of this equation, if $K \geq 6$. Therefore, we should simplify it.

We change

$$e^{-(\gamma'')^n \alpha} \approx 1 - (\gamma'')^n \alpha$$

and substitute it to the previous expression:

$$\sum_{n=2}^K (S_n - x_n) (n-1) = \alpha \sum_{n=2}^K S_n (n-1) (\gamma'')^{n-1}.$$

First of all we construct an approximated solution of this equation. In case when the meaning γ' is close to 1, we can use Taylor's row expansion:

$$(\gamma')^n = [1 - (1 - \gamma')]^n = 1 - n(1 - \gamma') + \dots$$

Applying this expansion we get

$$\sum_{n=2}^K (S_n - x_n) (n-1) \approx \\ \approx \alpha \sum_{n=2}^K S_n (n-1) [1 - (n-1)(1 - \gamma')].$$

Now we can estimate γ' :

$$\hat{\gamma}'' = 1 - \frac{\sum_{n=2}^K [S_n (\alpha - 1) + x_n] (n-1)}{\alpha \sum_{n=2}^K S_n (n-1)^2}.$$

If γ' is strongly different from the unit, this formula should be used for the first approximation to estimate.

For estimating the γ' parameter it is necessary to find an average number Θ of times when the words were not memorized and its mathematical expectation $M\Theta$. Θ can be calculated by experimental data:

$$\Theta = \frac{1}{N} \sum_{n=1}^K \sum_{i=1}^N r_i^n.$$

The mathematical expectation can be presented in the following way:

$$M\Theta = \frac{1}{N} \sum_{n=1}^K \sum_{i=1}^N M \xi_i^n = \frac{1}{N} \sum_{n=1}^K \sum_{i=1}^N M p_i^n.$$

Since $p_i^n = 1 - e^{-\alpha_i^n}$, then

$$M\Theta \approx \frac{1}{N} \sum_{n=1}^K \sum_{i=1}^N M \alpha_i^n.$$

Using the elements of probability theory and the method of mathematical induction we get the following expression (the k -th moment of the forgetting rate):

$$M(\alpha_i^n)^k = (\gamma')^k M(\alpha_i^{n-1})^k + \\ + [(\gamma'')^k - (\gamma')^k] M(\alpha_i^{n-1})^{k+1}.$$

According to this formula all the moments can be calculated. Further we can find $M\alpha_i^n$ and MQ .

$$M\alpha_i^{n+1} = M\{\gamma' \alpha_i^n e^{-\alpha_i^n} + \gamma'' \alpha_i^n (1 - e^{-\alpha_i^n})\} \approx \\ \approx \gamma' M\alpha_i^n + (\gamma'' - \gamma') M(\alpha_i^n)^2.$$

The estimate of γ' can be obtained from equality $M \approx \Theta$ and by construction of successive approximations, supposing that $M\alpha_i^1 = \alpha_i^1$ ($i = 1, 2, \dots, N$). The approximations should be constructed until $n = k$. Then we get $\hat{\gamma}'_{k-1}$ and take it as a parameter $\hat{\gamma}'$ estimate.

This approach essentially differs from usual training systems that as a matter of fact simulate behavior of a teacher along with his didactic godsend and failures. Suggested approach reduces an influence of a teacher and teaches a trainer, taking into account the individual features of memory. At every step it minimizes a distance between a learner's state and given purpose. It means that the training is optimal at every step, but it does not guarantee optimality of whole training process. We only get a solution close to optimal.

We took one of the approaches to intellectual technologies synthesis in training, developing new methods of training, proposed by professor L. A. Rastrigin [1]. It considers training process as the learner's memory state. But one should notice that this method is oriented for the properties of the concrete subject sphere vocabulary, i. e. is based on the multilingual frequency dictionaries.

References

1. Rastrigin, L. A. Contemporary Principles of Control of Complex Systems / L. A. Rastrigin. M.: Mir, 1985.
2. Rastrigin, L. A. Demo-Version «ASOLIJA-SistAn» / L. A. Rastrigin [et al]. Riga: Asolija, 1997.
3. Rastrigin, L. A. Adaptive Training With Trainer Model / L. A. Rastrigin, M. H. Erenstein. Riga: Zinatne, 1988.
4. Learner's Memory State Model in Informational Training Technology / H.-R. Fluck, M. V. Karaseva, T. A. Kovaleva and E. A. Souzdaleva // Siberian Region: abstr. of Interr. conf. Vol. 1. Krasnoyarsk, 1999. P. 3–10.
5. Kovalev, I. V. English-German-Russian Frequency Dictionary on System Analysis / I. V. Kovalev, M. V. Karaseva; Siberian State Aerospace University. Krasnoyarsk, 2004.
6. Kovalev, I. V. Adaptive algorithm for foreign vocabulary training on the basis of lexically related components / I. V. Kovalev, M. V. Karaseva, O. V. Leskov // Control systems and information technology. 2008. № 4 (34). P. 78–82.

М. В. Карасева

ИНФОРМАЦИОННО-ОБУЧАЮЩАЯ ТЕХНОЛОГИЯ: СОСТОЯНИЕ ПАМЯТИ МОДЕЛИ ОБУЧАЕМОГО

Рассмотрена проблема обучения иностранному языку, связанная с использованием модели состояния памяти обучаемого в информационно-обучающей технологии. Применен один из подходов синтеза интеллектуаль-

ных технологий в обучении, предложенный профессором Л. А. Растригиным. В качестве модели обучения была взята экспоненциальная зависимость вероятности незнания лексической единицы от скорости и времени заывания.

Ключевые слова: информационно-обучающая технология, состояние памяти модели обучаемого, процесс обучения, порция информации, алгоритм обучения.

© Karaseva M. V., 2009

УДК 681.3.006

А. С. Кузнецов

МОДЕЛИРОВАНИЕ РАСПОЗНАВАТЕЛЕЙ МУЛЬТИСИНТАКСИЧЕСКИХ ЯЗЫКОВ ПРОГРАММИРОВАНИЯ МУЛЬТИВЕРСИОННЫХ СИСТЕМ

Рассмотрено синтезированное абстрактное устройство распознавания мультисинтаксических языков программирования. Это устройство может быть использовано в качестве основы для разработки трансляторов языков программирования мультиверсионных программных систем.

Ключевые слова: мультисинтаксический язык, язык программирования, мультиверсионный подход, программные системы.

Для создания отказоустойчивых программных систем известно множество подходов, использующихся на практике. К их числу относится методология мультиверсионного, или N -вариантного программирования [1], одной из отличительных черт которой является разработка нескольких версий одного и того же функционального модуля или компонента. В целом данная методология является дополнением к таким способам получения отказоустойчивых систем, как тестирование и верификация программного обеспечения.

Мультиверсионная программная методология основывается на допустимом разнообразии входных данных, и ее можно ввести в следующие элементы процесса разработки программного обеспечения:

- языки программирования модулей;
- алгоритмы решения типовых задач;
- средства программирования;
- методы тестирования.

Авторами ряда работ по мультиверсионному подходу отмечается тот факт, что при разработке программного обеспечения важно применение различных языков программирования для оформления вариантов [2] или версий [3] программных модулей, поскольку реализация различных версий модуля на разных языках позволяет уменьшить вероятность появления в версиях однотипных ошибок, связанных с языком.

Процесс создания мультиверсионных программных систем подразумевает их архитектурное разбиение на компоненты и модули, создаваемые независимо друг от друга. Однако нередко возникает необходимость в кодировании этих модулей на разных языках в рамках одного и того же исходного текста, как, например при создании динамических HTML-страниц с включением кода на язы-

ках JavaScript и (или) VB Script [4]. Традиционный процесс разработки мультиверсионных систем не позволяет выполнить это простым способом. Обойти данное ограничение можно с помощью мультисинтаксических языков программирования [5] и автоматизированной системы MuYacc построения трансляторов языков этого класса.

Итак, в рамках одного исходного текста программы могут одновременно использоваться языки со свойствами, различающимися с точки зрения синтаксиса, и, как следствие, генерируемые разными классами грамматик. В связи с этим такие языки можно отнести к классу мультисинтаксических языков (МСЯ).

При этом синтаксис каждого из составляющих языков описывается контекстно-свободной грамматикой. Из языка-лидера выделяются специальные символы, сигнализирующие главному синтаксическому анализатору о передаче управления соответствующему вспомогательному анализатору. Количество лексических и семантических анализаторов может быть равно количеству используемых языков, а их взаимодействие друг с другом осуществляется по одной из традиционных схем.

Мультиавтомат с магазинной памятью. В основу системы MuYacc положено устройство распознавания МСЯ в виде мультиавтомата с магазинной памятью (МП-мультиавтомат). При этом каждый вспомогательный анализатор основывается на обычном автомате с магазинной памятью (МП-автомат), который принято описывать семеркой элементов [6]:

$$P = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F),$$

где Q – множество состояний автомата; Σ – алфавит входных символов; Γ – алфавит магазинных символов; δ – графически и таблично заданная функция переходов;