A NEW APPROACH TO COMPUTERIZED ADAPTIVE TESTING

KURAVSKY L.S.*, *Moscow State University of Psychology & Education, Moscow, Russia,*
e-mail: l.s.kuravsky@gmail.com
ARTEMENKOV S.L.**, *Moscow State University of Psychology & Education, Moscow, Russia,*
e-mail: slart@inbox.ru
YURYEV G.A.,*** *Moscow State University of Psychology & Education, Moscow, Russia,*
e-mail: g.a.yuryev@gmail.com
GRIGORENKO E.L.*,*** *Lomonosov Moscow State University, Moscow, Russia, Yale Universlty, New-Haven, USA,*
e-mail: elena.grigorenko@yale.edu

A new approach to computerized adaptive testing is presented on the basis of discrete-state discrete-time Markov processes. This approach is based on an extension of the G. Rasch model used in the Item Response Theory (IRT) and has decisive advantages over the adaptive IRT testing. This approach has a number of competitive advantages: takes into account all the observed history of performing test items that includes the distribution of successful and unsuccessful item solutions; incorporates time spent on performing test items; forecasts results in the future behavior of the subjects; allows for self-learning and changing subject abilities during a testing procedure; contains easily available model identification procedure based on simply accessible observation data. Markov processes and the adaptive transitions between the items remain hidden for the subjects who have access to the items only and do not know all the intrinsic mathematical details of a testing procedure. The developed model of adaptive testing is easily generalized for the case of polytomous items and multidimensional items and model structures.

Keywords: Markov processes, adaptive testing, IRT, computerized adaptive testing.

Introduction

Testing procedures are increasingly used in many contemporary applications requiring assessment of people or machine's behavior. According to conventional models of testing based on classical test theory for measuring the examinee's level in a specific skill or ability as precisely

For citation:
Kuravsky L.S., Artemenkov S.L., Yuryev G.A., Grigorenko E.L. A New approach to computerized adaptive testing. *Eksperimental'naya psikhologiya = Experimental psychology (Russia)*, 2017, vol. 10, no. 3, pp. 33—45. doi:10.17759/exppsy.2017100303

* *Kuravsky L.S.* Professor, Doctor in Technical Sciences, Dean of the Department of Information Technologies, Moscow State University of Psychology & Education, E-mail: l.s.kuravsky@gmail.com
** *Artemenkov S.L.* PhD, Professor, Moscow State University of Psychology & Education, E-mail: slart@inbox.ru
*** *Yuryev G.A.* Ph.D. in Physics and Matematics, associate professor, Deputy Dean of the Department of Information Technologies, Moscow State University of Psychology & Education, E-mail: g.a.yuryev@gmail.com
**** *Grigorenko E.L.* Professor, Head of Laboratory of Behavoir Genetics, Chair of Psychology of Education and Pedagogics, Lomonosov Moscow State University, Professor, Yale University, University of Houston, E-mail: elena.grigorenko@yale.edu

as possible these procedures usually should implement a big number of items that makes testing difficult to use. The way out is provided by computerized adaptive testing (CAT) that is greatly aided by the advent of modern technology and computing capabilities, emerged alongside the development of the Item Response Theory (IRT).

Using IRT or other CAT approaches is a method of test administration and latent constructs' measurements through as few testing items and as precisely as possible [8]. These constructs may include: abilities, attitudes, knowledge, skills, traits and other relevant categories. The term "construct" will be used below instead of any of these categories.

During a computerized adaptive testing procedure an *adaptive principle* of item selection is used, according to which the difficulties of the items required to implement must be in correspondence with the estimations of the subjects' attainment levels. According to IRT, this approach yields the best differentiation of subjects by their attainment level. The other advantage is represented by multidimensional computerized adaptive testing (MCAT), also including shorter testing time and a more accurate and efficient construct estimation. MCAT joins together theoretical and practical advancements that capitalizes on CAT, allows for access to more constructs of interest without adding the burden of additional pools of items to the instrument and eventually has increased preciseness (i.e., low standard error of measurement) [9]. Despite of all these benefits, available CAT and MCAT methods based on IRT are quite complex to implement and do not accommodate a number of special parameters as time, testing history, etc.

However, the principal problem associated with the IRT-based CAT and MCAT (which are usually combined with maximum likelihood evaluations) is approximate equality of probabilities for wrong and right solutions since a selected item difficulty must fit the subjects' attainment level estimation. This fact makes testing results dependent mainly from extraneous random factors which are not related to the constructs under study, thereby devaluing obtained conclusions that become dummy. Therefore, development of new reasonable approaches to CAT and MCAT is obviously urgent poser.

Presented below is one of the approaches to overcome it. The employed model is represented by *discrete-state discrete-time Markov processes (Markov chains)* to produce test items and obtain target estimations. A feature of the given approach is detection of item difficulties using limit state probability distributions obtained via model matrices of transition probabilities. Competitive advantages of the presented approach over the adaptive testing based on IRT (G. Rasch model) are as follows:

1. The estimation is not derived from local comparisons of current assessment evaluations and difficulties using the G. Rasch model but takes into account:

— All the observed history of performing test items, which includes the distribution of successful and unsuccessful item solutions and their order.

— Time spent on performing test items.

2. The assessments are based on forecasting results in the future provided that testing time is unlimited and they do not use local (i.e. for a certain item) comparisons based on the G. Rasch model, which may be unstable[1].

3. The number of items to be performed is substantially less.

[1] When abilities are approximately equal to difficulties in the classical logit scale, 0.5-logit shift in abilities yields 0.2-shift in probabilities. On the contrary, in case of the same condition, shift at 0.1 in the probability measure yields 0.25-logit shift in the ability measure. Thereby, the evaluations are quite sensitive to errors.

4. Selected item difficulties are linked to the history of performing test items and do not depend directly on the current estimations of subjects' attainment levels.

5. There is possibility for changing the difficulty of subject's constructs during a testing procedure owing to tiredness and other reasons.

6. Possibility of self-learning that results in improving the characteristics of adaptive testing model during its period of exploitation.

7. Easily available model identification procedure based on simple and rather accessible observation data.

At the same time, the presented approach and adaptive model under study may be considered as an extension of IRT, since the G. Rasch model is used as its element.

Advantages of the presented approach over the adaptive testing based on continuous-time Markov models are:

1. Discrete-time Markov processes are in use instead of the continuous-time ones [4—7], however the time spent on performing the items is taken into account via a time limit in the "trap" structures used in the adaptive model under study.

2. Polytomous items are allowed, with this possibility being easily implemented.

Model of Adaptive Testing

General Description of the Model

Utilized are *discrete-state discrete-time Markov processes* (*Markov chains*), with *transition probabilities* between their states being model parameters.

To describe how the probabilities of being in the given states are changed over time, Markov chains are applied. The typical structure presented by a scheme in Figure 1 is a finite chain with $2n+2$ states, in which transitions from state x_i ($i \neq 0$, $i \neq n$) are possible solely to the next state x_{i+1} or state x_{i*}. Available from states x_0 and x_n are only states x_1, x_{0*} and x_{n*}, respectively. Being in state x_{i*} ($i=0,...,n$) one can go to state x_i only.

States x_i and x_{i*} correspond to the i^{th} *substantial level of item difficulties*. A specific set of items of relevant difficulty content is defined for every *i*. As captured in Figure 1, states with a larger number include the items corresponding to higher difficulties than the states with a smaller number, with the "highest" level of difficulties corresponding to the rightmost state.

It is assumed that each subject has one of the specified *attainment levels* with indices $l \in \{0, ..., z\}$, where $(z + 1)$ is the number of these levels, with a set of items of a certain difficulty being assigned to each of these substantive levels and $z < n$. (Of note is that this particular illustration focuses on subjects' characteristics — i.e., constructs). At all attainment levels items are assigned to each substantive level of knowledge, abilities or skills.

Each attainment level corresponds to the certain interval of item difficulties containing more than one state (see Figure 1). Therefore, the number of difficulty levels is equal or greater than the number of attainment levels. The greater the attainment levels the higher the evaluation score.

As time functions, probabilities of being in model states are defined by the following matrix equation:

$$\boldsymbol{p}(t+1) = \mathbf{M}(\boldsymbol{\lambda}_l)\mathbf{p}(t),$$

where t is discrete time; $0 \leq t \leq T; t, T \in N$; T is the terminal time point; N is the set of natural numbers; $\mathbf{p}(t) = (p_0(t), \dots, p_n(t), p_{0*}(t), \dots, p_{n*}(t))^T$ represents probabilities of the being in model states in time point t; $\mathbf{M}(\boldsymbol{\lambda}_l) = \|m_{ij}(\boldsymbol{\lambda}_l)\|$ is the stochastic square matrix of transition probabilities between the Markov chain states, in which $m_{ij}(\boldsymbol{\lambda}_l)$ is the probability of the transition from state j to state i;

$\boldsymbol{\lambda}_l = (p^+_{0,l}, \dots, p^+_{n-1,l}, q^+_{0,l}, \dots, q^+_{n,l}, q^-_{0,l}, \dots, q^-_{n,l}, r_{0,l}, \dots, r_{n,l}, r_{0*,l}, \dots, r_{n*,l})^T$ is the ordered set of the transition probabilities in question for subject attainment level *l*. Square matrices have order $2n + 2$.

Classification of subjects is performed as it is presented in Section 2.

If a measurement scale is *continuous*, the entire range of its values should be divided into several intervals, each of which is interpreted as the certain substantive level of a state (i.e., constructs) under consideration. It is the interval of scale values that is to be selected as a result of the testing procedure. The greater the number of states, the more accurate this estimation. The more accurate the estimation, the greater the quantity of empirical data to be requested.

A testing procedure is ascertained by administering the items, the successful accomplishment of which requires specified constructs in the case of a certain attainment level. Difficulty of an item assigned to a subject corresponds to the model state occupied by him/her at the current time.

Below, the person who takes the test is referred to as *a subject*.

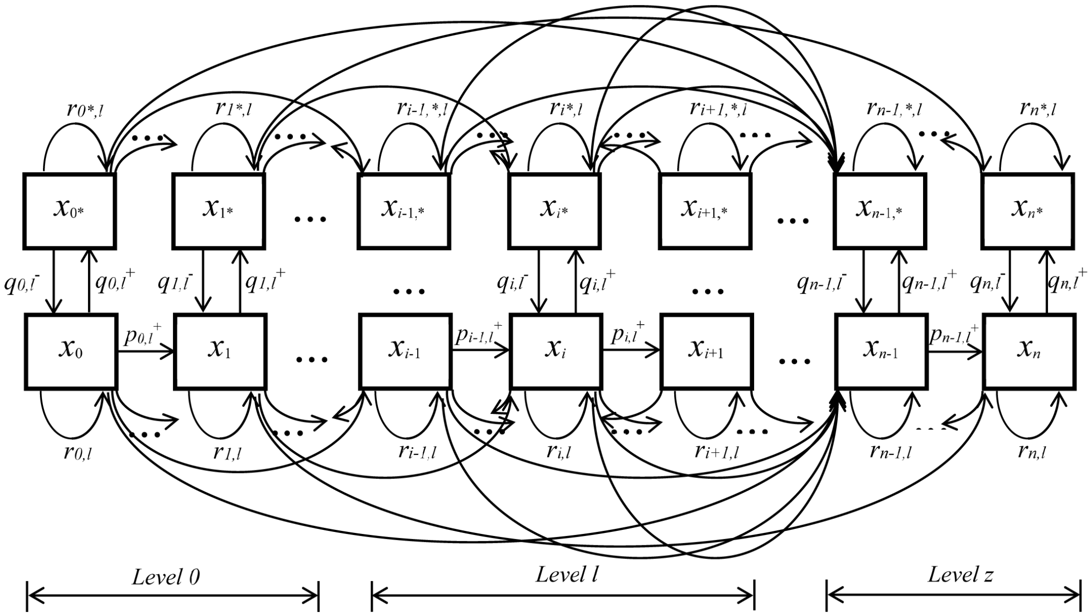

Figure 1. Discrete-time Markov chain representing a testing procedure composed of items: $\{x_i\}_{i=0,\dots,n}$ and $\{x_{i*}\}_{i=0,\dots,n}$ are the states, $\boldsymbol{\lambda}_l = (p^+_{0,l}, \dots, p^+_{n-1,l}, q^+_{0,l}, \dots, q^+_{n,l}, q^-_{0,l}, \dots, q^-_{n,l}, r_{0,l}, \dots, r_{n,l}, r_{0*,l}, \dots, r_{n*,l})^T$ is the set of transition probabilities between these states, $l \in \{0, \dots, z\}$ is an attainment level index.

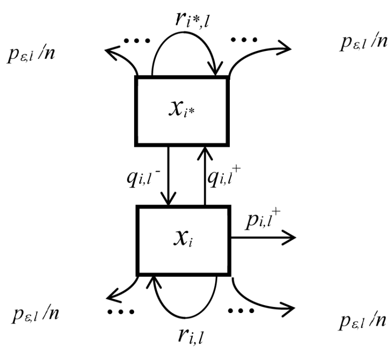

Figure 2. Element of the discrete-time Markov chain represented a testing procedure: "trap".

Transition probabilities are calculated the following way ($i = 0, \ldots, n$):

$q^-_{i,l} = p(i, l)k(t^*_i, i, l)(1 - p_{\varepsilon,l})$,

$r_{i,l} = (1 - p(i, l))\, k(t^*_i, i, l)(1 - p_{\varepsilon,l})$,

$r_{i*,l} = ((1 - p(i, l))\, k(t^*_i, i, l) + (1 - k(t^*_i, i, l)))\,(1 - p_{\varepsilon,l})$,

$q^+_i = (1 - k(t^*_i, i, l)(1 - p_{\varepsilon,l})$,

$$p(i, l) = \frac{e^{a(f(l) - i)}}{1 + e^{a(f(l) - i)}},$$

with $p^+_{i,l}$ being determined for $i = 0, \ldots, n - 1$:

$p^+_{i,l} = p(i, l)k(t^*_i, i, l)\,(1 - p_{\varepsilon,l})$,

where the last function expresses the G. Rasch dependence of probability $p(i, l)$ for performing a test item successfully on attainment level l and item difficulty level i; factors $k(t^*_i, i, l)$ represent probabilities for non-exceeding the corresponding time limits t^*_i in case of given difficulty levels i and attainment level l; is "error" probability for each state belonging to attainment level l; both function $f(l)$ and quantity a are the G. Rasch model parameters. Small probabilities $p_{\varepsilon,l}$ corresponding to the attainment levels are necessary to support testing adaptiveness as these values are responsible for transitions which are carried out in case of changing current estimations of attainment levels. They are assumed to be distributed uniformly over corresponding transitions from each state and, therefore, are equal to $p_{\varepsilon,l} / n$. Inclusion small "error" probability in the model is necessary to allow formally transitions, which are unlikely for a given attainment level but are characteristic for some other attainment level, otherwise these transitions would be impossible. This model element is useful for the subject classification. Small transition probabilities are averaged since their values are of the order of sampling error (despite the fact that they can be estimated using experimental data).

Appearance of "traps" shown in Figure 2 in the model is caused by both the necessity of taking into account time dynamics of a testing procedure (time is introduced in the model implicitly) and the possibility of differentiating easily two important groups of subjects: those who generate quickly a series of incorrect item solutions, and those who find a correct item solution for a long time period. This is essential for psychological diagnostics since this feature makes it possible to select people who are not critical with regard to results that they create. When $t^*_i \to \infty$, corresponding "traps" are transformed from 2-state to 1-state structures.

Transition Rules and Classification

At the initial time point of a testing procedure, a subject is assumed to be in model state x_0, i.e. the easiest item (i.e., capturing the lowest levels of construct) is administered.

When a subject is in state x_k the item assigned to him/her is selected randomly from the item set corresponding to the given state, with time limits t^*_i applied for each pair of Markov chain states (x_i, x_{i*}).

Subject's transitions between the states are determined by the following rules:

— If being in state x_i a subject performs the assigned item correctly and testing time does not exceed the prescribed value t^*_i (i.e., completes the assigned item within the specified time frame), he/she transits into state x_{i+1}.

— If being in state x_i a subject performs the assigned item incorrectly and testing time does not exceed the prescribed value t^*_i, he/she remains in state x_i.

— If a subject is being in state x_i and subject's time of performing the assigned item exceeds the prescribed value t_i^*, the subject transits into state x_{i^*}.

— If a subject is being in state x_{i^*} and subject's time of performing the assigned item either exceeds the prescribed value t_i^* or the subject performs the assigned item incorrectly and testing time does not exceed the prescribed value t_i^*, he/she remains in state x_{i^*}.

— If being in state x_{i^*} a subject performs the assigned item correctly and testing time does not exceed the prescribed value t_i^*, he/she returns into state x_i.

In fact, testing time is measured in attempts of performing items.

After each of the given transitions, state rectification via classification is carried out.

Classification is implemented to determine a subject's attainment level. Applied is the approach based on calculating the ultimate stationary distribution $\mathbf{p}^{\infty,l} = (p_0^{\infty,l}, \dots, p_n^{\infty,l}, p_{0^*}^{\infty,l}, \dots, p_{n^*}^{\infty,l})^T$ of the considered Markov chain state probabilities, which satisfies the following equation:

$$\mathbf{p}^{\infty,l} = \mathbf{M}(\boldsymbol{\lambda}_l)\mathbf{p}^{\infty,l}.$$

Distribution $\mathbf{p}^{\infty,l}$ is proved to be an eigenvector of currently selected stochastic matrix $\mathbf{M}(\boldsymbol{\lambda}_l)$, which exists always and corresponds to the unit eigenvalue. In case of matrices $\mathbf{M}(\boldsymbol{\lambda}_l)$ under consideration, it was also found that process $\mathbf{p}(t)$ is ergodic, other eigenvalues of the given matrix are strictly inside the unit circle in the complex plane), and $\mathbf{p}^{\infty,l} = \lim_{t\to\infty} \mathbf{p}^l(t)$. Numerical methods of calculating $\mathbf{p}^{\infty,l}$ are rather simple and can be found in reference books [1; 10].

To take into account subject trajectories between the states of the Markov chains under consideration and to make classification on the basis of distribution $\mathbf{p}^{\infty,l}$ more reliable and more fitting real situation, columns of matrices $\mathbf{M}(\boldsymbol{\lambda}_l)$, where $l \in \{0, \dots, z\}$, are corrected after each attempt to perform an item, viz.: in case of the transition between the model states corresponding to indices j and i of this matrices (from state j to state i) element m_{ij} is replaced by 1 and other elements of column j are replaced by 0, with these changes remaining valid only for a testing procedure of a given subject who implements them. Matrix $\mathbf{M}(\boldsymbol{\lambda}_l)$ in which such changes have been implemented is called *the matrix of a passed trajectory*. After every step of a testing procedure, the matrices of passed trajectories are calculated for each attainment level under study. Thereby, they keep information about the testing dynamics.

Using matrices of passed trajectories, mathematical expectation for index of ultimate pair of Markov chain states (x_e, x_{e^*}) is calculated to forecast subject attainment level l after each attempt to perform an item:

$$e_l = \sum_{i=0}^{n} i\,(p_i^{\infty,l} + p_{i^*}^{\infty,l}).$$

Before a testing procedure, stationary level difficulty index $e_{\infty,l}$ is to be determined for further evaluations for each attainment level l under consideration and matrix $\mathbf{M}(\boldsymbol{\lambda}_l)$ used at the initial time point of a testing procedure. To select the attainment level that fits the testing process under study best of all, one should calculate absolute differences between the obtained index e_l and stationary level difficulty indices $e_{\infty,l}$ for each attainment level l under consideration (see Figure 1) and then select minimum e_{min}:

$$e_{min} = \min_{l \in \{0, \dots, z\}} \left| e_l - e_{\infty,l} \right|.$$

Level l_{min} corresponding to this e_{min} value yields the best attainment estimation. It is obvious that the greater the subject correspondences to a given attainment level, the smaller the e_{min} values, and vice versa.

If level l_{min} differs from the current one, *state clarification* is implemented with the aid of transition to the state which corresponds to level l_{min} and has an integer index that is closest to the appropriate mathematical expectation e_l. State clarification is not carried out if either it requires the state number decrease after the right item solution or state number increase after the wrong item solution or exceeding the prescribed time limit. It is also expedient to perform these state number shifts if only they are not greater than a certain prescribed shift threshold.

If some matrix of a passed trajectory repeats itself at the previous testing procedure steps then the state number decrease should be executed. The shift value in use can be changed over time.

Matrix $(\boldsymbol{\lambda}_l)$ for the given attainment level l is called *consistent* if the mathematical expectation referred to above falls within the set of model states corresponding to this level (see Figure 1).

Upon being confronted with the procedure described here, a subject is "captured" in one of the states that fits best his/her assessment level of construct.

As an alternative way, *Bayesian estimations* can be used for classification. Knowing the model state in which a tested subject turns out to be after solving the last item at a specified time point, and the probability of being in this state at the specified time for each attainment level, which can be calculated using the previously given matrix equation, it is possible to estimate posterior probabilities of attainment levels with the aid of *the Bayes formulae*:

$$P(C_l \mid S) = \frac{P(C_l)P(S \mid C_l)}{\sum_{k=0}^{z} P(C_k)\, P(S \mid C_k)},$$

where C_l is an event indicating that a subject has reached the l^{th} attainment level ($l \in \{0, ..., z\}$), is an event indicating that a subject is located in the specified model state corresponding to a specified item difficulty level at the specified time, $P(C_l)$ is a prior probability for a subject to have reached the l^{th} attainment level, $P(S \mid C_l)$ is the probability of being in the specified model state at the specified time given that a subject has the l^{th} attainment level, and $P(C_l \mid S)$ is the probability of reaching the l^{th} attainment level given that a subject is located in the specified model state at the specified time.

The attainment level at which the highest conditional probability $P(C_{max} \mid S) = \max_l \{P(C_l \mid S)\}_{l=0,\dots,z}$ is reached yields the required classification. Probability distribution $\{P(C_l \mid S)\}_{l=0,\dots,z}$ is obtained as a result of performing the assigned item sequence, making it possible to estimate the reliability of the derived classification.

Model Identification

Identification of the Markov models under study is carried out using data samples consisting of subjects' testing outcomes. Each attainment level $l \in \{0, ..., z\}$ is processed separately and has its own identified matrix $\mathbf{M}(\boldsymbol{\lambda}_l)$, with a unique set of estimates of model parameters $\boldsymbol{\lambda}_l$ being associated with it. It enables further classification by figuring out the attainment level that has the best fit to a testing procedure under assessment.

In contrast to the approach presented in the papers [4—7], identification of the model under study does not require solving difficult optimization problems. If a relevant data base with observation results is available, it is reduced to the rather simple estimation of the following model parameters: function $f(i)$, value a, factor $k(t_i^*, i, d)$, and time limits $\{t_i^*\}_{i=0, \dots, n}$. This data base must make it possible to estimate frequencies of right and wrong solutions for each combination of subject attainment and item difficulty levels, with the prescribed time limit excesses being taken into account.

Parameters $f(i)$ and a can be estimated by *the maximum likelihood method* using some empirical data representing testing results for subjects with pre-defined attainment levels and items with pre-defined difficulty levels.

Factors $k(t_i^*, i, d)$ is determined directly with the aid of empirical data via the ratio of subjects who exceed and not exceed the corresponding time limits t_i^* in case of a given attainment level.

Parameters $\{t_i^*\}_{i=0, \dots, n}$ as well as “error” probability levels $p_{\varepsilon,l}$ can be determined by solving the optimization problem with criterion $C = \sum_{l=0}^{z} (e_l - e_{mean,l})^2$ to be minimized. Thereby, the time limit parameters are selected to make the expected ultimate pair of Markov chain states closer to a center of the model attainment level interval under consideration. Since actual value ranges for parameters $\{t_i^*\}_{i=0, \dots, n}$ are known in advance, the numerical method [3] of optimization problem solution can be used.

The Markov chain (see Figure 1) is identified separately for each attainment level.

Model Self-learning

When a certain attainment level is determined with the given accuracy C_* and after each attempt to carry out an item, probabilities m_{ij} corresponding to every already implemented transition between the states of this level model are replaced by the slightly increased values $m_{ij}(1 + \delta)$, where $\delta << 1$, with other elements of column j being reduced to the same small value $m_{ij}\delta / (2n + 1)$ so that the total sum of this column elements being kept equal to 1 (i.e. matrix $\mathbf{M}(\boldsymbol{\lambda}_l)$ being kept stochastic); herewith the presented matrix changes are saved for all subjects who carry out a given testing procedure during some defined observation period.

Presented series of small corrections for matrix elements actually implement the Kohonen self-learning method [2].

Features of the Testing Procedure

The testing procedure is terminated when one of the following events is the case:

— Characteristic value e_{min} becomes smaller than a certain prescribed *threshold value* (this case usually reduces the time of a testing procedure since this condition may be satisfied after a few items).

— Overall allotted procedure time limit is exceeded.

— The item assigned in state x_n is performed successfully without reaching the prescribed value t_n^*.

In the case of a *continuous* measurement scale, the approach presented can be used in the *“microscope” mode*, in which we get a rough estimate at the first stage using rather rough intervals. We then divide the stage interval where a subject finds himself at the end of a testing procedure into several subintervals of the smaller size, then repeat the testing procedure using the new Markov chain fitted for more accurate estimation with these subintervals on the second stage, and so on. The greater the number of these stages, the more accurate the estimation.

Both Markov chains and the abovementioned adaptive transitions remain hidden for the subjects who have access to unassigned items only and do not know all the intrinsic mathematical details of a testing procedure.

Polytomous Items

The model under study can be generalized for the case of *polytomous items*. Let us consider w possible variants of estimating the performance results for an item given in state x_i. In this case the corresponding model transition to state x_{i+1} that is shown in Figure 1 is replaced by a multivariate transition presented in Figure 3 (each state x_{i+1} has its own "trap"). Polytomous transition probabilities $\{p_0^+, \ldots, p_{w-1}^+\}$ are assumed to be proportional to empirical transition frequencies available via observations.

Item Difficulty

To estimate items' difficulties the relevant "duality theory" will be developed, in which subject's construct (e.g. ability) and item difficulty scales are considered as dual concepts replac-

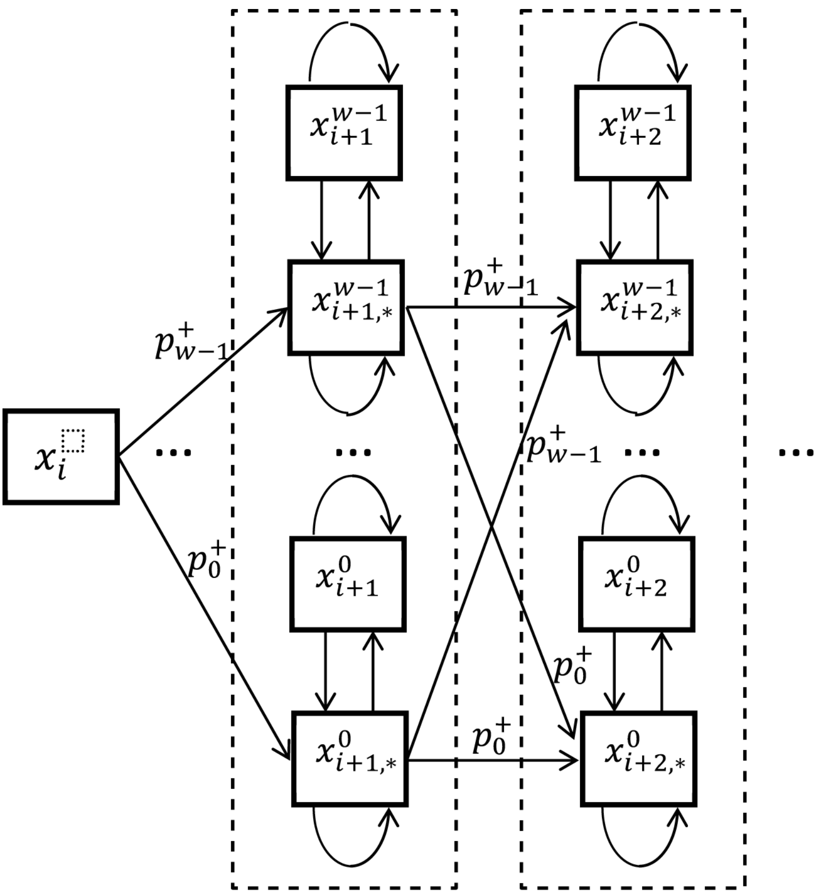

Figure 3. Transitions for polytomous items.

ing each other. The Markov chain that is used to estimate attainment levels can be applied to calculate difficulty levels if the states in question represent attainment levels instead of substantial levels in the previous model. The item to be evaluated "walks" along the Markov chain and is presented to the subjects which attainment level corresponds to the state where the item under study find itself at the given moment. All the theory features including the model structure, transition rules, model identification as well as classification are kept the same.

Multidimensional Items

When a testing procedure contains items evaluated using several measurement scales, the aforementioned assessment for each scale can be performed independently using the model presented in Figure 1. The results obtained may be represented with the aid of multidimensional structures composed of "traps" shown in Figure 2. For example, 3-D structure of this type is given in Figure 4 (its constituent elements are abovementioned "traps").

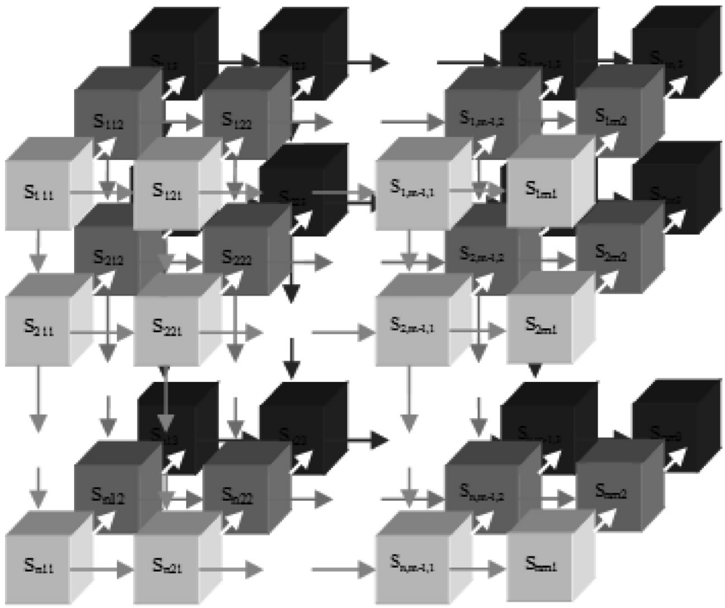

Figure 4. 3-D structure (constituent elements are "traps": see Figure 2)

Program for demonstration

The program demonstrating features of the adaptive testing model under consideration is available at URL http://it.mgppu.ru/files/model.zip.

Results and Conclusions

A new approach to adaptive testing is presented on the basis of discrete-state discrete-time Markov processes. This approach is based on an extension of the G. Rasch model used in IRT and has decisive advantages over the adaptive IRT testing. Its principal features are:

1. The model of adaptive testing takes into account all the observed history of performing test items that includes the distribution of successful and unsuccessful item solutions.

2. The developed model incorporates time spent on performing test items; it is taken into account via a time limit in the "trap" structures used in the adaptive model.

3. The assessments in the model are based on forecasting results in the future behavior of the subjects and allow for changing the value of subject's constructs during a testing procedure and self-learning that results in improving the characteristics of adaptive testing model during its period of exploitation.

4. Discrete-time Markov chain processes used instead of the continuous-time ones facilitate the adaptive testing model that implies easily available identification procedure based on simply accessible observation data.

5. Selected item difficulties are linked to the history of performing test items and do not depend directly on the current estimations of subjects' attainment levels.

6. The developed model of adaptive testing is easily generalized for the case of polytomous items and multidimensional model items and structures.

7. The proposed adaptive approach is made ready for CAT and MCAT implementation.

Funding
This work is supported by Grant 17-06-00277 from the Russian Foundation for Basic Research.

References

1. Burden R.L., Faires J.D., Numerical Analysis, Brooks/Cole, Cengage Learning, 9th Ed., 2011. 895 pp.
2. Kohonen T. Self-Organizing Maps, Springer, 3th Ed., 2001. 501 pp.
3. Kuravsky L.S., Marmalyuk P.A., Yuryev G.A. and Dumin P.N. A Numerical Technique for the Identification of Discrete-State Continuous-Time Markov Models// Applied Mathematical Sciences. Vol. 9, 2015, No. 8, pp. 379—391. URL: http://dx.doi.org/10.12988/ams. 2015.410882.
4. Kuravsky L.S., Margolis A.A., Marmalyuk P.A., Panfilova A.S., Yuryev G.A., Dumin P.N. A Probabilistic Model of Adaptive Training. — Applied Mathematical Sciences, Vol. 10, 2016, no. 48, 2369—2380, http://dx.doi.org/10.12988/ams.2016.65168.
5. Kuravsky L.S., Margolis A.A., Marmalyuk P.A., Yuryev G.A., Dumin P.N. Trained Markov Models to Optimize the Order of Tasks in Psychological Testing. — Neirokomp'yutery: Razrabotka, Primenenie (Neurocomputers: Development, Application), 2013, No. 4, pp. 28—38 (in Russian).
6. Kuravsky L.S., Marmalyuk P.A., Baranov S.N., Alkhimov V.I., Yuryev G.A. and Artyukhina S.V. A New Technique for Testing Professional Skills and Competencies and Examples of its Practical Applications. — Applied Mathematical Sciences, Vol. 9, 2015, no. 21, 1003—1026, http://dx.doi.org/10.12988/ams.2015.411899.
7. Kuravsky L.S., Yuryev G.A. Applying Markov Models to Processing Testing Results. — Voprosy Psychologii (Problems of Psychology), 2011, No. 2, pp. 112—121 (in Russian).
8. Thompson N.A., Weiss D.J. A framework for the development of computerized adaptive tests. Practical Assessment, Research & Evaluation, 2011, 16(1), 1—9.
9. Torre J. de la, Patz R.J. Making the Most of What We Have: A Practical Application of Multidimensional Item Response Theory in Test Scoring. Journal of Educational and Behavioral Statistics, 2005, 30(3), 295—311. doi:10.3102/10769986030003295.
10. Wilkinson J.H. The Algebraic Eigenvalue Problem, Oxford, Clarendon Press, 1988. 662 pp.