УДК 681.3

M. V. Karaseva

MULTILINGUAL TECHNOLOGY. SYSTEM ASPECTS OF ITS ORGANIZATION AND USAGE

System aspects of multilingual adaptive-training technology organization and usage to train multilingual vocabulary are considered. Information vocabulary support of the technology is presented. Its realization is directed at the architectural layer of rules primary, corresponded to the functions of adaptive-training model and multilingual term basis.

Keywords: multilingual approach, adaptive model, system architecture, frequency term dictionary, adaptive-training technology.

Adaptive-training technology is developed as a model approach of the training problem, proposed and developed by professor L. Rastrigin [1]. The relation between a teacher and a learner is considered as the relation between a controlled and controlling units. It permits to use the control theory methods. It's difficult to construct the precise model, that is why it is necessary to construct the approximate model and to adapt its parameters corresponding to a real unit (a learner). This approach, realized for multilingual adaptive-training technology is considered. The main applied aspect of the technology is a practical problem solution of the foreign vocabulary study and intensive store of the profession-oriented vocabulary for specialists, students, working with foreign literature or listening to lectures in foreign languages.

Multilingual adaptive-training technology. One of the ways of foreign profession-oriented vocabulary study is to use the algorithm of training, operated on the basis of the adaptive model of a learner, taking into account the individual peculiarities of memory and forgetting [2].

Formally one can present memory as a totality of a great number of cells. Let the cell number of the native words be

$$K_N = \{k_1, \ldots, k_N\},$$

where each cell contains a native word.

When one studies the first foreign language (e. g. English), a new number of cell to memorize new English words is generalized as a result of the supplement of new words:

$$K_E = \{k_1, \ldots, k_E\},$$

Further, associations of an i-native word (e. g. Russian) appear with its English equivalent

$$A_i^{NE}, \quad i = 1, \ldots, N.$$

When one studies the second foreign language (e. g. German) in a period of time, new cells appear to memorize German words.

Let's denote a number of cells as

$$K_G = \{k_1, \ldots, k_G\}.$$

A set K_G is a bit higher then a set of memorized English words, as a number of elements K_E reduces depending on the time passed after the previous memorizing.

One can see the dependence of time and memory capacity in fig. 1. Newly memorized German words take up former «English cells», which are empty by the time. The information is transformed from short-term into long-term memory.

New associations of native words with their German equivalents appear as well:

$$A_i^{NG}, \quad i = 1, \ldots, N.$$

The degree of their relation varies depending on the time of memorizing and forgetting.

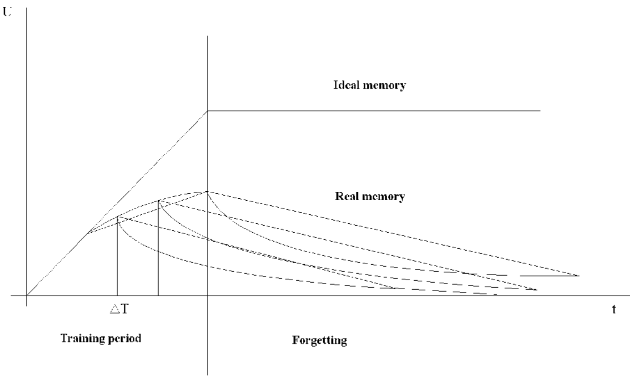

Fig. 1. Variation of the total learned information capacity while memorizing and forgetting

Using the information multilingual adaptive-training technology word meanings of all simultaneously learned languages store in the same cell and fixed there at the same time. And the time passed after memorizing is the same. Forgetting rates could be a bit different because of the subjective personal peculiarities, but all the rest factors will be identical for both words [2].

So called multiple associations appear. Let's take as an example tripartite associations with English and German equivalents:

$$A_i^{NEG}, \quad i = 1, \ldots, N.$$

The call process to a set of cells happens only once and associations don't conflict.

The associated parameter expressing the constraint degree of an i-native word with its two foreign equivalents can be presented as:

$$F_i(A^{NEG}) = 1 - e^{-v_0 n}, \tag{1}$$

where n – number of seances; v_0 – information perception rate:

$$v_0 = \frac{M_n}{\lambda T_n}, \tag{2}$$

where M_n – memorizing capacity U_n; T_n – n-seance duration; $0 < 1 < 1$ – loss information coefficient while memorizing.

The main components of the information-algorithmic support of the multilingual training technology are computer systems, realizing adaptive training algorithm and electron frequency term dictionaries.

The main advantages of this computer system are mentioned below:

– word frequency is taken into account (to memorize words more frequently used in the text);

– individual peculiarities of memorizing are granting;

– seance intervals are unrestricted;

– information amount in comparison with the total information amount is taken into account;

– multilingual approach provides associative field generation round memorized terms;

– forgetting rate reduction of the memorized information while repetition is taken into account.

System aspects of multilingual technology organization. The basis of any methodology is a system architecture as well as certain strategies, analysis and design methods. The architecture of modern systems is three-layered and it has the following characteristics:

– strictly defined layers;

– formal and explicit interfaces between layers;

– invisible and protected details inside a layer.

According to the mentioned requirements system architecture of the multilingual adaptive-training technology is presented in fig. 2.

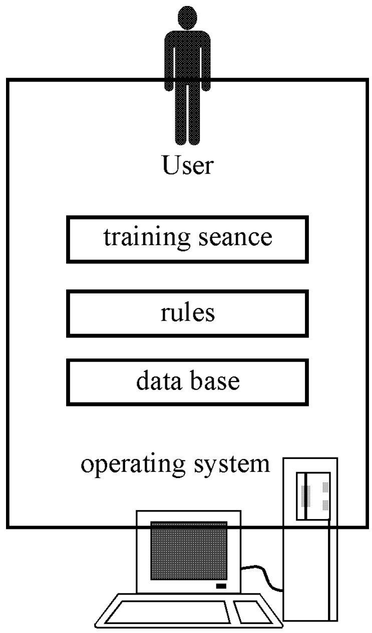

Fig. 2. System architecture of multilingual technology

The three layers (data base, rules, a seance) reflect the increase of the abstraction level of the system architecture. The most detailed level is data base, more elevated abstraction level is a level of rules and the most elevated abstraction level is a level of sйances [3]. This architecture contains a layer of rules which is a relatively new concept directed toward a multilingual approach and corresponded to adaptive-training model functions.

The table 1 present three-layer system architecture of the multilingual technology according to the modern structure methodology and its steps of development (requirement analysis, design, realization).

So multilingual adaptive-training technology is regulated by structural system methodology and adapted for three-components architecture owing to the priority of the rule layer, corresponding to the functions of adaptive-training technology within the bounds of multilingual term basis.

Information – term support of the multilingual technology. To form information-term basis of the multilingual-training technology it is necessary first of all to develop electronic frequency multilingual dictionaries for different fields. These dictionaries are the term basis for the developing computer system to study foreign vocabulary. Term selection and term frequency effect the training system. The pointed characteristics are included into the training model and influence the quality of training process.

So, term set, corresponding to the basis information component of the multilingual training technology can be described the following way:

Multilingual component = {language term 1, language term 2, …, language term *N*, language frequency 1, language frequency 2, …, language frequency *N*}/

DSSD uses an analogous notation (fig. 3).

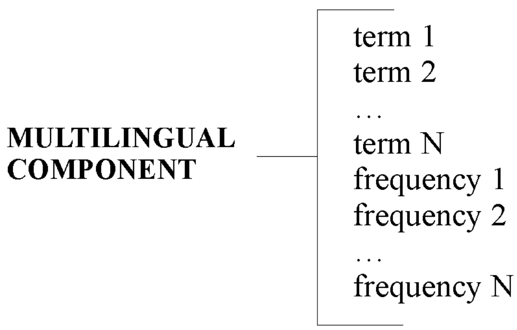

Fig. 3. Description of the basis information component using DSSD notations

Constructs of Varnie–Orra diagrams are obvious from the example (fig. 4). Two basis constructs of Varne–Orra diagrams are given: hierarchy and sequence [2]. They can be interpreted the following way: the first stage of the multilingual technology is a seance structure selection of multilingual components; the second stage is determination of characteristics and requirements; the third stage is training.

Table 1

System stages of three-layered architecture of the multilingual technology

Layers	Analysis	Design	Realization
Seance of training	Stream of seances	Stream of interactive interaction	Dialogue «user-system»
Rules	Stream of processes	Model of components	Program
Data base	Data model	Data base scheme	Multilingual term tables, electronic dictionaries etc.

Logic of data processing for multilingual technology, defined by Varnie–Orra diagrams and supported by program and algorithmic means is described below.

Multilingual technology usage while developing the system of vocabulary training. We modified the approach, proposed by professor Z. A. Rastrigin. In this case a learner is an object of control, and a teacher or trained device is a source of control [1].

Look at the training process scheme (fig. 5). Y is a trainer's state, measured by data unit; Y' – information about trainer's state got by a teacher in answer to questions U, besides U includes the portion of training information. Purpose of training Z^* and resources R are given to the teacher.

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

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

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

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

$$p_i^n = p_i(t_i^n) = 1 - e^{-} e^{\alpha_i^n t_i^n}, \qquad (3)$$

where α_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:

$$\alpha_i^{n+1} = \begin{cases} \alpha_i^n, \text{if } i \notin U_n, \\ \gamma' \alpha_i^n F_n^{\text{NEG}}, \text{ if } i \in U_n \text{ and } r_i^n = 1, \\ \gamma'' \alpha_i^n F_n^{\text{NEG}}, \text{ if } i \in U_n \text{ and } r_i^n = 0{,}5, \\ \gamma''' \alpha_i^n F_n^{\text{NEG}}, \text{ if } i \in U_n \text{ and } r_i^n = 0, \ n = 1, 2, ..., \end{cases} \qquad (4)$$

where α_i^1 – initial value of the rate of forgetting, estimated by the maximal probability method according to the expression (6), $0 < \alpha_i^1 < 1$, $(i = 1, 2, ..., N)$; $\hat{\alpha} = -\ln \frac{M_n - x}{M_n}$, x – a number of elements unremembered from M_n; γ', γ'', γ''' – parameters defining the individual features of trainer memory, estimated by the maximal probability method, $0 < \gamma' < 1$, $0 < \gamma'' < 1$, $0 < \gamma''' < 1$.

An answer to tests $R_n = (r_1^n, ..., r_{Mn}^n)$ can be written as:

$$r_i^n = \begin{cases} 1, \text{ if there is no answer}, \\ 0{,}5, \text{ if there is no answer at least in one language}, \\ 0, \text{ if answer at right}. \end{cases} \qquad (5)$$

The effectiveness criterion Q_n 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:

$$Q_n = \sum_{i=1}^{N} p_i(t_i^n) q_i \to \min, \qquad (6)$$

where $p_i^n(t_i^n)$ – probability of i-th TI element ignorance; q_i – a relative frequency of the element appearance in the text, $0 < q_i < 1$:

$$q_i = \frac{q_i^{\max}}{V},$$

where $q_i^{\max} = \max q\{q_{i1}, q_{i2}, q_{i3}\}$ – an absolute frequency of the element appearance in the text, q_{i1}, q_{i2}, q_{i3} – frequencies of English, German and Russian word from the corresponding multilingual frequency dictionary; V – text volume for frequency dictionary.

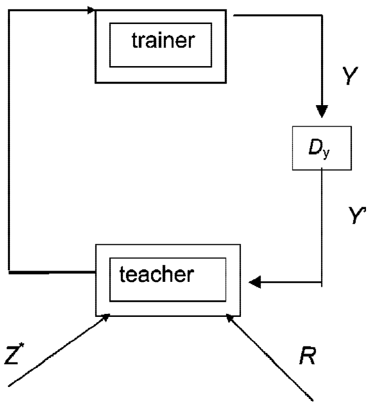

Fig. 5. Training process scheme

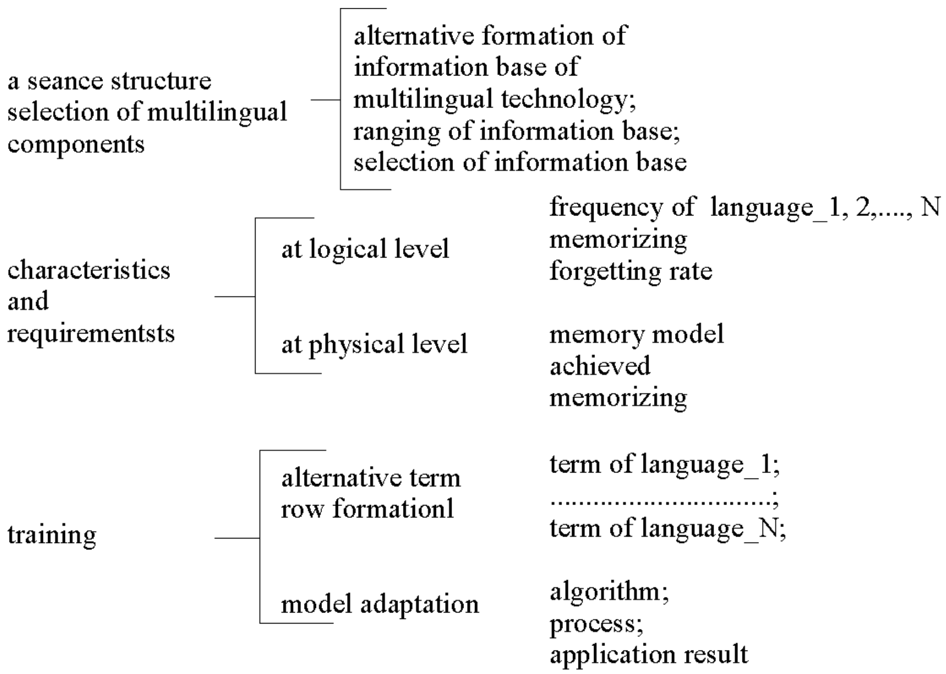

Fig. 4. Varne–Orra diagram for the main stages of multilingual technology.

To minimize Q_n it is necessary to include the elements of the greatest meaning of multiplication $p_i(t_i^n)q_i$ 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:

$$u_1 = \arg \max_{1\le i\le N} p_i(t_i^n)q_i^{\max},$$

$$u_2 = \arg \max_{1\le i\le N,\, i\ne N} p_i(t_i^n)q_i^{\max}, \qquad (7)$$

$$u_{M_n} = \arg \max_{1\le i\le N} p_i(t_i^n)q_i^{\max},$$

$$i \ne u_j \ (j = 1, 2, \ldots, M_n - 1),$$

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

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

$$M_n = \max_{1\le M\le N} \left\{ M : T_n \ge k \sum_{i\in\{u_1,\ldots,u_M\}} p_i(t_i^n) \right\}, \qquad (8)$$

where k – average learning time of TI element at its first presentation to the trainer; $u_1, \ldots, u_M$ – numbers of TI elements. Parameter k is unknown a priori and should be adapted:

$$k_{n+1} = k_n + \mu\,(T_n' - T_n), \qquad (9)$$

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

The training is completed when Q_n is of the required level of training δ. A number of seances n determines a duration of training, when $Q_n \le \delta$.

So, one can distinguish the following stages of the training algorithm:

1. Check knowledge of portion TI and as a result construct a set of answers to the test according to the expression (5).
2. Realize parameter adaptation of the model according to the rule (5) taking into account expressions (1) and (2).
3. Correct the ignorance probability vector of TI elements, i. e. form $P_n + 1$ according to the expressions (3) и (4) and taking into account forgetting time after last learning tin:

$$t_i^{n+1} = \begin{cases} \Delta t_n, \text{ если } i \in U_n, \\ t_i^n + \Delta t_n, \text{ если } i \notin U_n, \ n = 0, 1, \ldots . \end{cases}$$

4. Calculate Q_{n+1} according to the expression (6).
5. If $Q_{n+1} \le \delta$, then the training process is terminated; otherwise you should define TI portion U_{n+1} by expression (7), and present this portion to the learner taking into account rules (8) and (9).

The authors compared proposal training modal with already known training models. The experiment of comparison is in the following:

– to make a plan of the experiment, sufficient to all models;
– to receive an experimental curve which points determine a part of incorrect answers at each stage of training;
– to estimate model's parameters by experimental data and theoretical training curve construction for each model;
– to estimate the proximity of the theoretical and experimental curves by the chosen criterion;
– choice of the model, describing the training process best.

For the experiment training models of Rastrigin, Bush–Mosteller, Miller–McHill, Terstone, Restle and Krichevski were chosen [1]. Theoretical training curves were constructed for all models and collected in one diagram. The proximity level of the theoretical training curve by the experimental one was estimated as:

$$\rho_i = \sum_{n=0}^{K-1} \left| \Theta(n) - M\Theta^{(i)}(n) \right|,$$

where i – model number, $i = 1, 2, \ldots, 7$; $K = 8$; $\Theta(n)$ – points of the experimental curve; $M\Theta^{(i)}(n)$ – points of theoretical training curve, received with a help of the i-th model.

The receive results (table 2) show that multilingual-adaptive model coincide rather well with experimental data, so points of the theoretical training curve are closed to the points of the experimental curve.

The training algorithm and information support has been realized in the multilingual computer system to train profession vocabulary «Virtual Teacher 1.0» (fig. 6), working under Windows 9x/2000.

The system is aimed at intensive accumulation of the foreign vocabulary and is an advanced tool for it. With a help of «Virtual Teacher 1.0» the first version of which is used in a practice the vocabulary of three languages can be learnt simultaneously.

The way of the work with the system «Virtual Teacher 1.0» is in fig. 7. It is possible to use frequency dictionaries of different fields, done according to the system format.

The system provides the opportunity to store user's data and to come back to an interrupted seance of training. A training coefficient is done in real time. Having achieved the required level a user is recommended to take the following seance of training.

One can update dictionaries, attach new dictionaries, form one's own dictionaries. It is possible to use simultaneously a separate program module with concordance functions.

Table 2

Results of the experiment

Model number	Model	ρ
1	Multilingual-adaptive	0,124
2	Adaptive	0,135
3	Bush–Mostller	0,406
4	Miller–Mchill	0,893
5	Terstone	0,458
6	Restle	0,160
7	Krichevski	0,697

So, a thoroughly selected complex of training programs (e. g. a text editor, Internet browser, trivial electronic dictionaries and encyclopedias) provides an opportunity to create computer training environment, as an integrated mean of foreign language study on the basis of the described system.

This paper mainly studies system aspects of a new multilingual adaptive-training system technology. The approach usage let use information basis, model and training algorithms more effectively. The approach enlarges professional vocabulary of a few foreign languages simultaneously.

The authors presented modification of training model and algorithm, using associative parameter, which shows the relation of a native word with its foreign equivalents. The training algorithm has been realized in the multilingual computer system «Virtual Teacher 1.0» to memorize professional vocabulary. A lexical basis is a special English-German-Russian frequency dictionary of System Analysis for students of this specialization as well as interpreters, specialists who deals with system analysis problems.

Bibliography

1. Rastrigin, L. A. Adaptive training with a puple as a model / L. A. Rastrigin, M. Kh. Erenshtein // Riga : Zinatne, 1988.

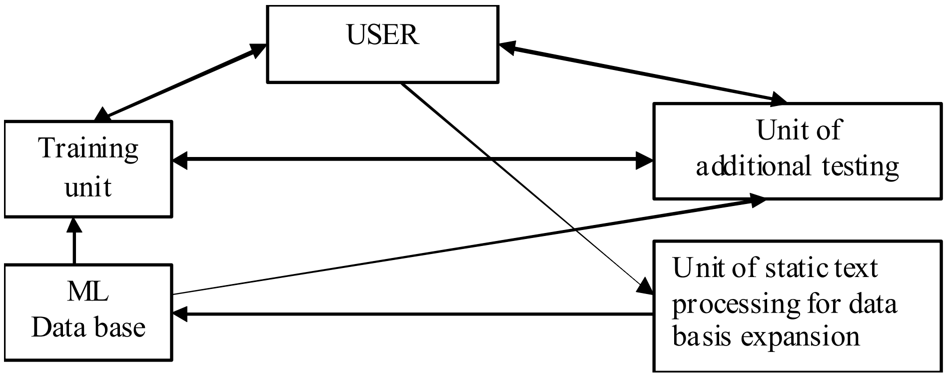

Fig. 6. Conceptual scheme of system

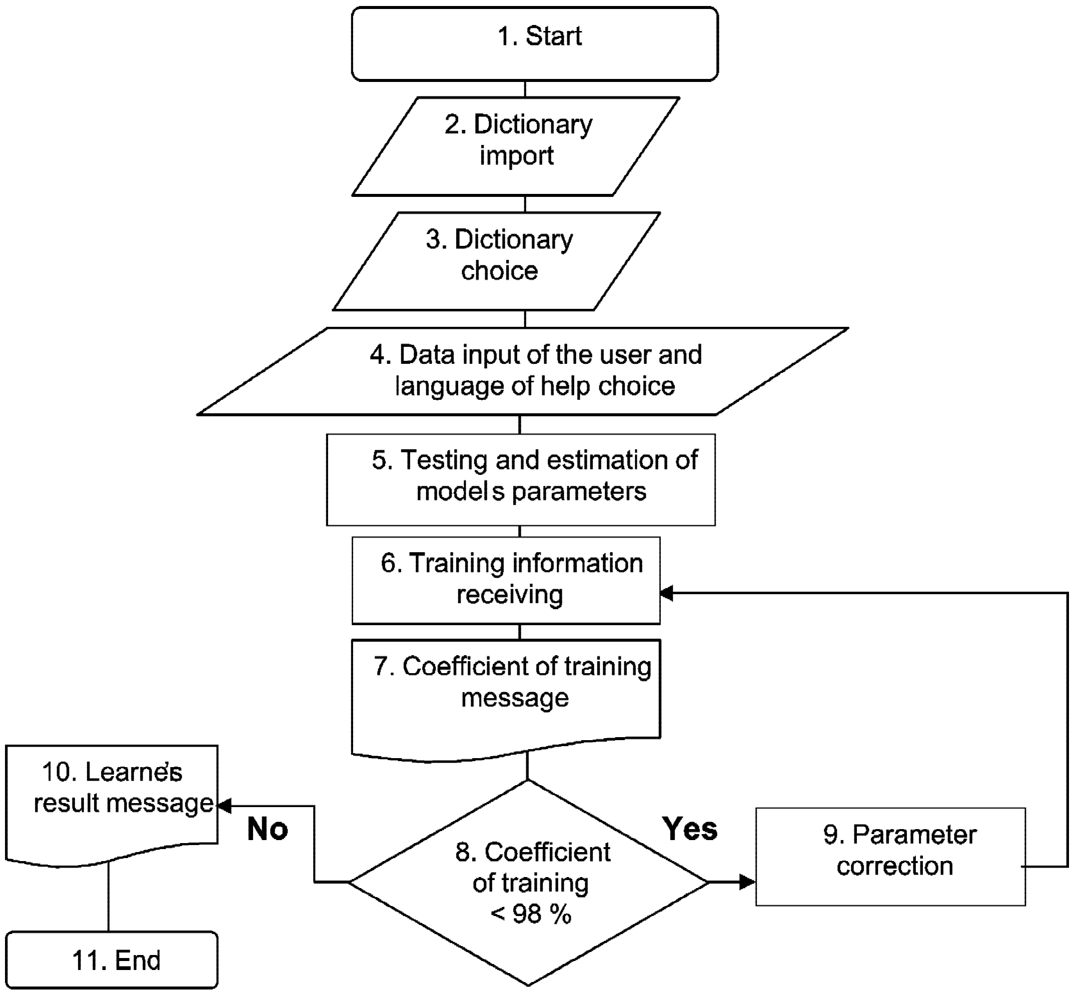

Fig. 7. System work scheme

2. Kovalev, I. V. Information training technology: learner's memory state model / I. V. Kovalev, M. V. Karaseva, E. A. Souzdaleva // Advances in Modeling : J. of AMSE Periodicals. Ser. B. 2002. Vol. 45, № 1. P. 21–32.

3. Kovalev, I. V. Adaptive algorithm of foreign vocabulary training on the basis of lexically related components / I. V. Kovalev, M. V. Karaseva, O. V. Leskov // Control systems and information technology : Sci. J. 2008. № 4 (34). P. 78–82.

M. В. Карасева

МУЛЬТИЛИНГВИСТИЧЕСКАЯ ТЕХНОЛОГИЯ. СИСТЕМНЫЕ АСПЕКТЫ ОРГАНИЗАЦИИ И ПРИМЕНЕНИЯ

Рассмотрены системные аспекты организации мультилингвистической адаптивно-обучающей технологии и ее применение для изучения многоязычной терминологической лексики. Представлено информационно-терминологическое обеспечение этой технологии, реализованное с учетом ориентации на первичность архитектурного слоя правил, соответствующего функциям адаптивной модели обучения и многоязычному терминологическому базису.

Ключевые слова: мультилингвистический подход, адаптивная модель, системная архитектура, частотный терминологический словарь, адаптивно-обучающая технология.

УДК 519.688

М. П. Варыгина, И. В. Киреев, О. В. Садовская, В. М. Садовский

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ АНАЛИЗА ВОЛНОВЫХ ДВИЖЕНИЙ В МОМЕНТНЫХ СРЕДАХ НА МНОГОПРОЦЕССОРНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ[1]

Для численного исследования динамических задач моментной теории упругости Коссера на многопроцессорных вычислительных системах разработаны параллельные алгоритмы, программная реализация которых выполнена по технологии SPMD на языке Fortran-95 с использованием библиотеки передачи сообщений MPI. Программный комплекс оснащен средствами сжатия больших массивов данных с контролируемой потерей информации, позволяющими многократно снизить сетевой трафик при копировании файлов – результатов счета с удаленного кластера и служащими для компактного хранения численных решений в постоянной памяти компьютера.

Ключевые слова: моментный континуум, упругие волны, высокопроизводительные вычисления.

Моментная теория упругости Коссера [1; 2] служит для описания деформации материалов с микроструктурой: композитов, гранулированных, порошкообразных, микроразрушенных и микрополярных сред. В отличие от обычной теории упругости в ней неявно присутствует малый параметр среды – характерный размер частиц микроструктуры, поэтому при численном решении задач расчеты необходимо выполнять на сетках, размер ячеек которых меньше этого параметра. При решении динамических задач в пространственной постановке эффективными оказываются параллельные алгоритмы, поскольку они позволяют распределять вычислительную нагрузку между большим числом узлов кластера, что дает возможность существенно измельчать расчетные сетки, повышая тем самым точность численного решения.

Математическая модель. В модели моментной среды, кроме поступательного движения, которое характеризуется вектором скорости v, рассматриваются независимые повороты частиц с вектором угловой скорости щ, и наряду с тензором напряжений у, компоненты которого несимметричны, вводится несимметричный тензор моментных напряжений m. Полную систему уравнений модели образуют уравнения движения, кинематические соотношения и обобщенный закон линейной теории упругости:

$$\begin{gathered}\rho\dot{v} = \nabla\cdot\sigma + \rho g,\\ j\dot{\omega} = \nabla\cdot m - 2\sigma^a + jq,\\ \dot{\Lambda} = \nabla v + \omega, \quad \dot{M} = \nabla\omega, \qquad (1)\\ \sigma = \lambda(\delta : \Lambda)\delta + 2\mu\Lambda^s + 2\alpha\Lambda^a,\\ m = \beta(\delta : M)\delta + 2\gamma M^s + 2\varepsilon M^a,\end{gathered}$$

[1] Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 08–01–00148), Комплексной программы фундаментальных исследований Президиума Российской академии наук № 2 «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация» и Междисциплинарного интеграционного проекта Сибирского отделения Российской академии наук № 40.