УДК 81’32

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДЛЯ НЕЗАВИСИМОГО КОМПЬЮТЕРНОГО ПРЕДСТАВЛЕНИЯ СЛОЖНЫХ ВЫРАЖЕНИЙ В ЕСТЕСТВЕННЫХ ЯЗЫКАХ

Б.Ж. Баячорова, П.С. Панков

Представлена методика независимого компьютерного представления основ естественных языков, приведены примеры условных выражений и подчиненных оборотов.

Ключевые слова: компьютерное представление языка; независимое представление языка; сложное выражение.

MATHEMATICAL MODELS FOR INDEPENDENT COMPUTER PRESENTATION OF COMPLEX EXPRESSIONS IN NATURAL LANGUAGES

B.J. Bayachorova, P.S. Pankov

The method of independent computer presentation of foundations of natural languages is presented; examples of models of conventional expressions and subordinate clauses are given.

Keywords: computer presentation of language; independent presentation of language; complex expression.

1. Introduction. We put the problem of completely independent presentations of natural languages (see definitions of underlined words below). Combining ideas of [1] and [2] we gave suggestions and have developed elements of such presentations [3], [4], [5], [6], have implemented a primary version of an algorithmic language for such purposes [7], proposed basic items for a language [8], and abstract notions [9].

To meet the common demands of testing (validity, objectivity and reliability) and to improve efficiency of testing we [10] used the following principles: generativity; uniqueness; concreteness.

The software for complex computer examination in Kyrgyz language based on such principles approved by the National commission on state language and is used in educational institutions in Kyrgyzstan.

The aim of this paper is to propose possible ways of independent presentation of some combinations of words of a natural language.

2. Definitions and Hypotheses. One of main tasks of present day informatics is developing of interactive computer presentations of all familiar real and virtual objects to offer the user the opportunity to master them safely and effectively before real treating.

Definiton 1. If a computer presentation of an object does not depend on the user’s knowledge and skills on similar objects then we call it independent.

In our opinion, such presentations are more effective because the user can learn inductively – without referencing other objects in mind. In regards with learning a language, the user begins to thinking in it, without translation in mind.

The following notions are used by us in computer testing.

Definiton 2. Generativity means that a complete task must not exist before the testing and must be generated (randomly) just in it [11].

Generativity is used also for presentation of notions with random auxiliary objects.

Definiton 3. Uniqueness means that all examinees must obtain different versions of tasks.

Definiton 4. Concreteness means that the user’s respond may be: number, word, short phrase, action.

The following definitions describe a language (both human and computer) from our standpoint.

Definition 5. If low energetic outer influences can cause sufficiently various reactions and changing of the inner state of the object (by means of inner energy of the object or of outer energy entering into object besides of commands) at any time then such (permanently unstable) object is an affectable object, or a subject, and such outer influences are commands.

Definition 6. A system of commands such that any subject can achieve desired efficiently various consequences from other one is a language.

Hypothesis 1. A human's genuine understanding of a text in a natural language can be clarified by means

of observing the human's actions in real life situations corresponding to the text.

Definition 7. Let any notion (word of a language) be given. If an algorithm acting at a computer: generates (randomly) a sufficiently large amount of instances covering all essential aspects of the notion to the user, gives a command involving this notion in each situation, perceives the user's actions and performs their results clearly on a display, detects whether a result fits the command, then such algorithm is said to be a computer interactive presentation of the notion.

Certainly, commands are to contain other words too. But these words must not give any definitions or explanations of the notion.

Definition 4. If all words being used in Definition 3 are unknown to the user nevertheless s/he is be able to fulfil the meant action (because it is the only natural one in this situation) then the notion (word of a language) is said to be primary. If the user has to know supplementary words to complete the action then the notion is said to be secondary. Thus, there is a natural hierarchy of notions.

Using this method we can present not only real notions (objects and actions) but also notions which have imaginary concepts. Also, some computer games can be adapted to develop an "adequate" representations of fictional and abstract concepts or objects [12].

Hypothesis 2. A person learning a natural language without references to any other ones, hearing a notion in various situations begins to form a kind of mathematical model in mind corresponding to this notion by means of trial and error method and attempts to fulfil operations similar to mathematical ones: closing and compactification. After successful completing such operations, the human feels "mastering" this notion.

Hypothesis 3. Any notion has a minimalistic model (involving minimal number of entities in Occam's sense).

Some notions mean fuzzy logic [13], we extended it to verbs [14], too.

3. Entities and their Programming. As all features of the existing world can be represented by a computer and by a language and there ought to be a closer correlation between them. We propose the programmer to implement such correlation by means of entities:

Time – is obvious and implemented as a sequence of actions and conditions of types *after*, *before*.

Subset, disconnected sets (primary verbs: PUT, TAKE, ENTER, GO OUT).

Plurality (CHOOSE, FIND, GATHER, BRAVE (the bravest of some persons); is implemented as an array.

1D (line) (PULL, LENGTH, BETWEEN) is implemented as a real number - coordinate.

2D (plane) (ENTER, GO OUT, AREA, BOUDARY, INTERIOR, EXTERIOR) is implemented as a pair of real numbers - coordinates.

Multilayer-2D (COVER, OPEN, SHIFT) is implemented as a triple of pair of real numbers – coordinates and a natural number of layer.

3D (space) (TIE, UNTIE, VOLUME, SURFACE) and implemented as a triple of real numbers – coordinates.

Animated (SHOW, GIVE, STRONG) is programmed as a stirring object.

Motion (STOP)

Person (ENTER, GO OUT, CLEVER, BRAVE, COWARD) is programmed as an avatar.

Attraction (including gravitation) and repelling (FALL, RISE, UP, DOWN, VERTICAL, HORISONTAL, HIGH) are implemented on vertical plane by motion (slow falling down), also up.

Size (LITTLE, BIG, THIN, THICK, SHORT, LONG).

Colors. Physically, they are reduced to differences in lengths of light waves but due to human's perceptions they are to be an entity.

Topological transformations (CUT, GLUE, CONNECT).

Sufficient transformation – includes physical and chemical ones.

Tale (imaginary world) – since references to worlds with other laws and transformations (tales, science fiction and mathematical objects) exist in the language they ought to be presented too.

4. Ways to Define Notions Non-Verbally. Some of the following ways imitate the way young children form notions: Naturalness (there exists the only natural action in this situation fitting the command); Possibility (the only object can be used to fulfil the action); Presence (both the object and the corresponding word appear simultaneously); Similarity (two objects have a same property and the corresponding word is repeated twice too); Using an etalon; Extrapolation (the user is to imagine the future or the past by observation of a video clip).

5. Some Types of Expressions.

5.1. Conventional expressions.

First step. "IF THE BALL IS RED THEN PUT IT INTO THE LEFT BOX!

IF THE BALL IS GREEN THEN PUT IT INTO THE RIGHT BOX!"

Red balls and green balls are appearing alternatively and randomly.

Second step. "IF THE BALL IS RED THEN PUT IT INTO THE LEFT BOX OTHERWISE PUT IT INTO THE RIGHT BOX!"

Balls of different colors are appearing one by one randomly.

Third step. “IF THE OBJECT (THING) IS A BALL THEN PUT IT INTO THE LEFT BOX OTHERWISE PUT IT INTO THE RIGHT BOX!”

5.2. Subordinate clauses (there are many balls in various positions on the screen). “PUT THE BALL WHICH IS UNDER THE TABLE INTO THE BOX!”

Other kinds of expressions in various languages can be presented in such a way.

References

1. *Asher J*. The strategy of total physical response: An application to learning Russian / J. Asher // International Review of Applied Linguistics. 1965. № 3.
2. *Winograd T*. Understanding Natural Language / T. Winograd. New York: Massachusetts Institute of Technology, 1972.
3. *Панков П.С.* Обучающая и контролирующая программа по словоизменению в кыргызском языке на ПЭВМ / П.С. Панков. Бишкек: Мектеп, 1992. 20 с.
4. *Pankov P.S*. Active English on computer / P.S. Pankov, J.Sh. Aidaraliyeva, V.S. Lopatkin // Conference "Improving Content and Approach in the Teaching of English Language in the Context of Educational Reform". Bishkek, 1996. P. 25–27.
5. *Pankov P.S*. Virtual Environment for Interactive Learning Languages / P.S. Pankov. Human Language Technologies as a Challenge for Computer Science and Linguistics // Proceedings of 2nd Language and Technology Conference. Poznan, 2005. P. 357–360.
6. *Pankov P.S*. Algorithmic Interactive Presentation of Notions [Текст] / P.S.Pankov, S.A.Alaeva, V.A. Kutsenko // Proceedings of the 11th International Conference "Speech and Computer" SPECOM'2006. St.-Petersburg, 2006. P. 478–480.
7. *Pankov P.S*. Algorithmical Language for Computer-Based Presentation of Notions / P.S. Pankov, P.S. Dolmatova // 4th International Conference on Electronics and Computer. Almaty: Suleyman Demirel University, 2007. P. 274–279.
8. *Bayachorova B.J.* Independent Computer Presentation of a Natural Language / B.J. Bayachorova, P.S. Pankov // Varia Informatica. Lublin: Polish Information Processing Society, 2009. P. 73–84.
9. *Bayachorova B.* Independent Computer Presentation of Abstract Notions of Natural Languages / B. Bayachorova, P. Pankov // Actual Problems of Computer Science. Lublin: ECCC Foundation, 2013, No. 1(3). P. 64–69.
10. 10. *Pankov P*. Software for Complex Examination on Natural Languages [Текст] / P. Pankov, P. Dolmatova. Human Language Technologies as a Challenge for Computer Science and Linguistics // Proceedings of 4th Language and Technology Conference. Poznan, 2009. P. 502–506.
11. *Панков П.С*. Опыт и перспективы использования комплекса UNIQTEST уникальных тестовых заданий в учебном процессе / П.С. Панков, Ж.Р. Джаналиева // Тез. докл. научно-практич. конф. «Образование и наука в новом геополитическом пространстве». Бишкек, 1995. С. 217.
12. *Borubaev A.A*. Spaces Uniformed by Coverings [Текст] / A.A. Borubaev, P.S. Pankov, A.A. Chekeev. Budapest: Hungarian-Kyrgyz Friendship Society, 2003 (Chapter 4. Constructive and computer presentations of uniform spaces).
13. *Zadeh L.A*. The concept of a linguistic variable and its application to approximate reasoning / L.A. Zadeh // Information Sciences, 1975. Vol. 8. P. 199–249, 301–357; Vol. 9. P. 43–80.
14. *Pankov P.S.* Mathematical Models for Independent Computer Presentation of Turkic Languages / P.S. Pankov, B.J. Bayachorova, M. Juraev // TWMS Journal of Pure and Applied Mathematics, 2012. Vol. 3, No. 1. P. 92–102.