I T H E A

International Journal

INFORMATION THEORIES & APPLICATIONS

2010 Volume 17 Number 4

International Journal
INFORMATION THEORIES & APPLICATIONS
Volume 17 / 2010, Number 4

International Journal "INFORMATION THEORIES & APPLICATIONS" (IJ ITA)
is official publisher of the scientific papers of the members of
the ITHEA International Scientific Society

IJ ITA welcomes scientific papers connected with any information theory or its application.
IJ ITA rules for preparing the manuscripts are compulsory.
The **rules for the papers** for IJ ITA as well as the **subscription fees** are given on *www.foibg.com/ijita*.

Responsibility for papers published in IJ ITA belongs to authors.

General Sponsor of IJ ITA is the **Consortium FOI Bulgaria** (www.foibg.com).

International Journal "INFORMATION THEORIES & APPLICATIONS" Vol. 17, Number 4, 2010

Printed in Bulgaria

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with:
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine,
Institute of Mathematics and Informatics, BAS, Bulgaria,
Universidad Politécnica de Madrid, Spain.

Publisher: **ITHEA**
Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, www.foibg.com, e-mail: info@foibg.com

ISSN 1310-0513 (printed) **ISSN 1313-0463 (online)** **ISSN 1313-0498 (CD/DVD)**

ORIENTATION ESTIMATION WITH APPLICATIONS TO IMAGE ANALYSIS AND REGISTRATION

David Asatryan, Karen Egiazarian, Vardan Kurkchiyan

Abstract: *In this paper, a novel method for image orientation angle determination is proposed. The method is based on the use of parameters of scattering ellipse of an image gradient field. Various problems with application of the proposed technique are considered. The results of the experiments for determining the predominant direction of an image are presented. Three methods of image orientation angle (OA) estimation are compared. It is shown that the method based on scattering ellipse technique is more precise than the histogram-based methods. The term of the "backlash" of image rotation algorithms is submitted and a method for "backlash" estimation is proposed. The "backlash" can be used for determination of the quality of arbitrary rotation algorithm. It is also shown that the proposed technique can be applied to the problems related to the Horizon line tracking.*

Keywords: *Image registration, orientation angle, directionality, regularity, gradient, rotation, scattering ellipse.*

ACM Classification Keywords: *Image Processing and Computer Vision*

Introduction

Image registration is the process of determining the correspondence between all the points in two images of the same scene. Registration of the images is requires in image analysis applications that involve two or more images of a scene. The reference and the referred image could be captured at different times, using different devices, from different angles etc.

The principal problem which is being solved by image registration methods is to find a geometric transformation such that the referred (or template) image becomes similar to the reference image. Successive solution of this problem depends on prior knowledge, available about possible type of deformations of referred image, which could be arrived due to the method of image capturing. Therefore there are huge numbers of image registration methods, described in the scientific literature. We can refer to [1-2], devoted to the survey on image registration methods and applications as well.

The most of registration procedures include image rotation operation, which is performed by the appropriate interpolation technique. The rotation can be performed in two ways. The first way is based on the scanning a set of possible values of rotation angle and performing the image rotation at each value. Then the image of maximal similarity with the reference image is considered as a referred one. The second way is based on preliminary estimation of the difference of orientation angles of the reference and referred images, and rotation the referred image at that angle.

There are a variety of techniques in the literature to estimate orientation angle, including methods using image gradients [3], directional histograms [4], discrete Radon transform [5, 6], modified Radon Transform [7] etc.

In this paper, the second way is considered, therefore a technique for image OA based on using scattering ellipse of image gradient field and related parameters is proposed. The gradient field is obtained by Sobel operator.
The proposed method is based on determination of the scattering ellipse of an image gradient field. The main axis of that ellipse is used for image predominant direction angle determination. In the present paper, the results of various experiments are given, which allows to compare different estimators of image rotation angle, some rotation methods and determinate the "backlash" of rotation algorithm.

Method

The method originates in a new approach to the problem of image quality assessment, described in [8]. According to that approach the measure for quality assessment is based on the usage of the image gradient magnitude distribution. Such measure gives the results similar to the human visual system. But the image gradient field also contains information on image orientation, notably the dominant direction (see, for example, [3]).
Let $I = \{I(m,n)\}$ be a Gray Scale image of size $M \times N$, where $I(m,n)$ is the pixel intensity with coordinates (m,n), $m = 0,1,...,M-1$, $n = 0,1,...,N-1$. Let's denote by matrixes $\|G_H(m,n)\|$ and $\|G_V(m,n)\|$ the horizontal and vertical gradients of an image at the point (m,n), by $\|F(m,n)\|$ the gradient magnitude, and by $\|A(m,n)\|$ the gradient angle, where

$$F(m,n) = \sqrt{G_H^2(m,n) + G_V^2(m,n)}\,, \tag{1}$$

$$\|A(m,n)\| = \left\| \operatorname{arctg} \frac{G_V(m,n)}{G_H(m,n)} \right\|, \tag{2}$$

Consider the gradient field as a two-dimensional sample from random variable with the elements

$$G(m,n) = [G_H(m,n), G_V(m,n)]. \tag{3}$$

Define the scattering ellipse of gradient field (3) as the solution of equation with respect to the variables (g_H, g_V) [9] as follows

$$\frac{1}{1-\rho_{HV}^2}\left[\frac{(g_H-\mu_H)^2}{\sigma_H^2} - \frac{2\rho_{HV}(g_H-\mu_H)(g_V-\mu_V)}{\sigma_H\sigma_V} + \frac{(g_V-\mu_V)^2}{\sigma_V^2}\right] = C^2, \tag{4}$$

where μ_H, μ_V, σ_H, σ_V are the sample mean and sample standard deviation of the horizontal and vertical components of the gradient field, ρ_{HV} is the correlation coefficient between them, C is a constant. The large axis of the ellipse coincided with the line of orthogonal regression, and the slope angle α is determined by formula as follows

$$\operatorname{tg}\alpha = \frac{2 * \rho_{HV}}{\sigma_H^2 - \sigma_V^2 + \sqrt{\left(\sigma_H^2 - \sigma_V^2\right)^2 + 4\rho_{HV}^2}}. \quad (5)$$

We assume that the angle α is measured from the horizontal axis anticlockwise as it is shown in Figure 1.

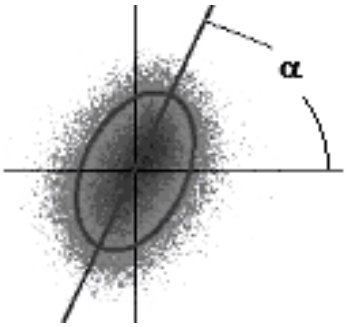


Figure 1. An image scattering ellipse and the orientation angle.

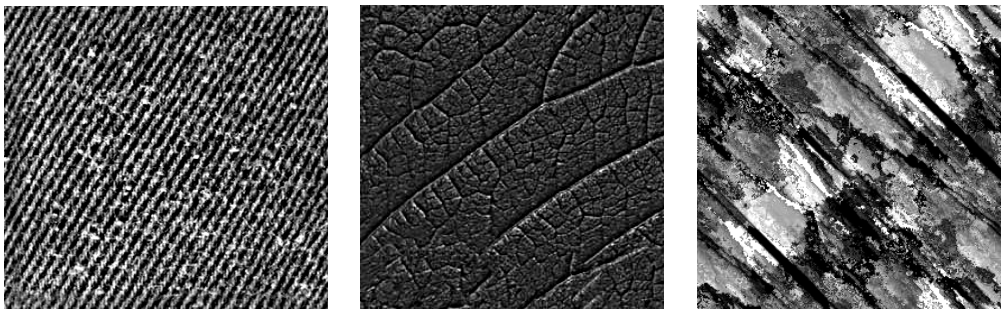
Image orientation angle determination

In this section the results of some experiments, using the technique above, are given and analyzed.

Let I be an original image, I_α be the image I rotated by an angle α using certain rotation algorithm.

Described method of image orientation angle estimation can be applied to certain problem, which is arisen during the registration process, such as:

- determination of an image orientation angle with respect to some prior fixed direction;
- determination of difference between orientation angles of two different parts of an image;
- determination of the rotation angle shift between two different images of the same scene;
- error analysis of different rotation algorithms;
- Horizon line tracking, etc.
-



$\alpha = 65°$ $\alpha = 18°$ $\alpha = 136°$

Figure 2. Determination of the predominant direction of the images.

Experiment 1. Determination of the predominant orientation angle of an image. Three images with predominant orientations, which could be distinctly estimated by visual analysis, are shown in Figure 2. Using the above described method we estimate the angle values which are given below the images. We can see that these values quite correspond to the visual estimation results.

Figure 3. Pisa Tower. The chosen parts of the image are separated by rectangles.

Experiment 2. Determination of the Pisa Tower leaning angle. It is known that the Pisa Tower is leaned at an angle of 5.5 degrees (see Figure 3). We can estimate this angle by subtraction of orientation angle $\hat{\alpha}_{\mathrm{T}}$ of the Tower and orientation angle $\hat{\alpha}_{\mathrm{T}}$ the building on the left of the Tower.

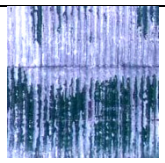
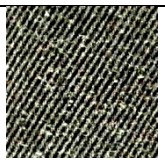
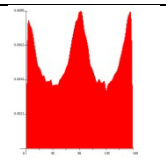
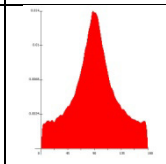
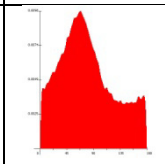
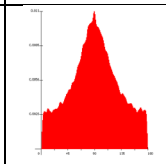
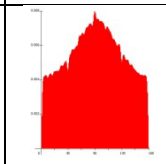
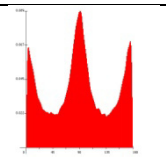
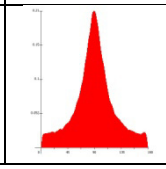
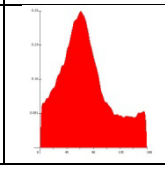
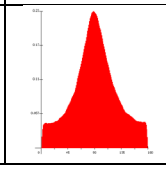
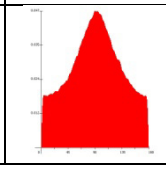
Appropriate areas of image are chosen to increase the estimation precision, as it is shown in Figure 3. The left part is of size 450x180 pixels, and the right part is of size 450x620 pixels. As a result of estimating by formula (5) we obtain $\hat{\alpha}_{\mathrm{B}} = 90.5°$, $\hat{\alpha}_{\mathrm{T}} = 85.1°$, $\hat{\alpha}_{\mathrm{B}} - \hat{\alpha}_{\mathrm{T}} = 5.4°$.

Comparison of orientation angle estimating methods

In this section we compare different methods for OA estimation. Image rotation is performed by algorithm, described in [10].

Experiment 3. Error determination for orientation estimation algorithm. Let $\hat{\alpha}_0$ be the orientation angle of original image I. Rotate the image I by an angle α and estimate the orientation angle $\hat{\alpha}_r$ of the image I_α. Angle $\hat{\alpha}_r - \hat{\alpha}_0$ can be interpreted as the error of orientation estimation algorithm. This approach allows comparing various estimation algorithms by values of mean-square error (MSE) at different rotation angles.

Table 1. Samples of texture images and histograms given by algorithms 2 and 3.

1					
2					
3					

In this experiment we compare three methods of orientation angle estimation:

- Method using the scattering ellipse orientation (Method 1);
- Method based on histogram of individual gradient angle (Method 2);
- Method based on histogram of the gradient magnitude, whereby for the angle α (Method 3), by formula as follows

$$H(\alpha) = \begin{cases} \sum_m \sum_n F(m,n), & \text{if } A(m,n) = \alpha \\ 0 & \text{otherwise} \end{cases},$$

where $A(m,n)$ is the gradient direction, determined by formula (5).

Table 1 shows samples of selected images and corresponding histograms (rows 2 and 3) determined by above specified algorithms. It can be noted that the algorithm 3 provides the histogram with less variation; therefore it can be considered as more precise estimation of the rotation angle.

Table 2. The mean square error of rotation angle estimation using different methods.

Angle α, degree	Estimation method		
	1	2	3
5	0.739	3.250	3.540
10	1.283	4.366	3.810
15	1.651	4.573	4.419
20	2.132	4.034	6.965
25	2.342	6.888	6.271
30	2.544	8.562	6.365
35	2.477	7.522	7.265
40	2.273	5.028	5.405
45	1.934	4.414	3.282

Rotation of 15 selected images by angles from 5º to 45º by discrete of 5º was performed. The results are given in Table 2. We can see that the algorithm 1 is the most precise and MSE is equal to 1-2 degrees, while other algorithms are worse.

Determination of the "backlash" of an image rotation algorithm

Experiment 4. Determination of the "backlash" for rotation algorithms. The term "backlash" we introduce to describe the results of application of various methods of image rotation, when at first the image is rotated by an angle α, and then it is rotated in backward direction by the same angle. In ideal case the last image will coincide with the original image. But because of distortions due to inevitable application of certain interpolation or other processing technique these images will differ. This difference we call "backlash" of rotating algorithm.

The "backlash" of estimation algorithm includes the following steps:

Step 1. Rotate image I by an angle α.

Step 2. Estimate the rotation angle by subtracting the orientation angles of images I_{α} and I. Let $\hat{\alpha}$ be the result of subtracting.

Step 3. Rotate the image I_{α} by an angle $-\alpha$. Let $\mathrm{I}_{\alpha,-\alpha}$ be the resulted image.

Step 4. Estimate the rotation angle of image $\mathrm{I}_{\alpha,-\alpha}$ by subtracting the orientation angles of images $\mathrm{I}_{\alpha,-\alpha}$ and I.

The result $\hat{\alpha}_b$ of subtracting is interpreted as the "backlash" of rotation algorithm.

We consider three algorithm of image rotation:

- Using B-spline interpolation technique by [10];
- Using Bi-linear interpolation technique;
- Using Cubic spline interpolation technique.

The MSE of "backlash" for each rotation algorithm applied to the same 15 images are collected in Table 3. We see that there are no significant differences between MSE values for these algorithms, in spite of the differences of used interpolation technique. We also can note that the MSE values don't depend on rotation angle.

Table 3. MSE of "backlash" for various rotation algorithms.

Angle $\hat{\alpha}_b$, **degree**	**Algorithm**		
	1	**2**	**3**
5	0.418	0.592	0.696
10	0.439	0.618	0.706
15	0.442	0.639	0.658
20	0.441	0.623	0.638
25	0.439	0.661	0.650
30	0.438	0.632	0.712
35	0.439	0.644	0.661
40	0.439	0.642	0.804
45	0.441	0.623	0.673

Experiment 5. Determination of the "backlash" for rotation angle estimation algorithm. Here the same 15 images above are rotated using the Algorithm 1 by different angles. Table 4 includes the MSE for specified estimation technique applied to "backlash" determination.

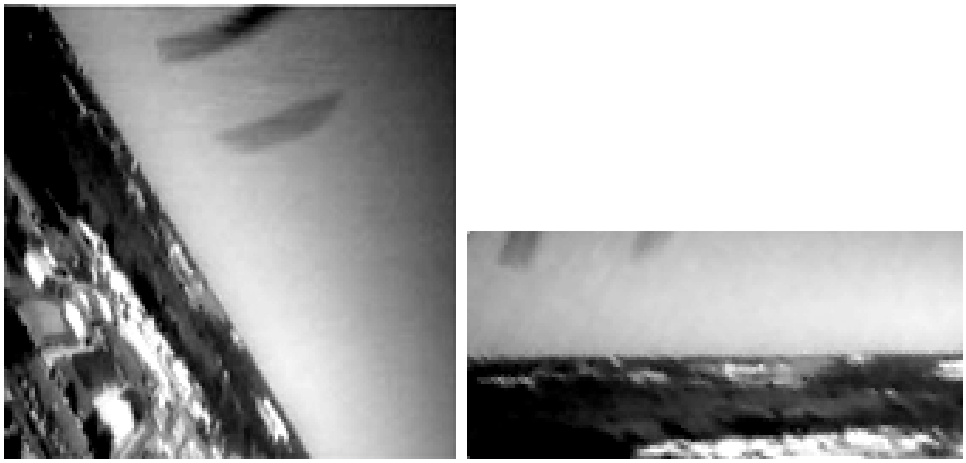
Table 4. MSE of the "backlash" for different estimation methods.

Angle α, **degree**	**Estimation method**		
	1	**2**	**3**
5	0.696	4.960	1.299
10	0.706	2.179	1.847
15	0.658	2.844	2.031
20	0.638	4.496	2.062
25	0.650	4.998	0.935
30	0.712	4.398	0.433
35	0.661	5.054	1.346
40	0.804	4.313	0.968
45	0.673	5.500	1.118

We see that the method based on scattering ellipse properties is better.

Horizon line tracking

Experiment 6. *Horizon line tracking* is an important task for certain applications of image processing methods, such as structural interpretation of seismic images, underground object tracking, coast line determination by remote sensing, flight control of small unmanned aircraft etc. In these cases the central problem is image orientation angle estimation. The knowledge of image orientation angle allows to rotate the image anticlockwise at that angle and to get the image of "normal" orientation, i.e. as if it was taken in the horizontal plane.
An image example containing the sky-ground border (or Horizon line) is shown in Figure 4a. The predominant direction of orientation is estimated by proposed method and resulted to ~60°. A fragment of the image after rotating by angle of 60° counterclockwise is shown in Figure 4b. The residual orientation angle of Figure 4b is equal about 1°.



a b

Figure 4. Image with sky-ground border taken at unknown viewpoint. a – original, b- a fragment of rotated image.

Conclusion

In this paper, a new technique for image orientation angle estimation is proposed. The idea is based on using the slope angle of the main axis of the scattering ellipse of image gradient field. By using the proposed technique some problems related to image orientation determination are considered and the results of numerical experiments are described. Two more methods of orientation angle estimation based on gradient angle histogram are compared with that. The effectiveness of the proposed method is shown. The term of image rotation algorithm "backlash" is submitted and a method for "backlash" estimation is proposed. The "backlash" method is used for determination of the quality of arbitrary rotation algorithm. It is also shown that the proposed technique can be applied to the problems related to horizon line tracking.

Bibliography

[1] Medha V. Wyawahare, Pradeep M. Patil, and Hemant K. Abhyankar, Image Registration Techniques: An overview. International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 2, No.3, pp 11-27, 2009.

[2] Zitova Barbara, Flusser Jan, Image registration methods: a survey, Image and Vision Computing, 21, pp 977-1000, 2003.

[3] Mester R., Orientation estimation: Conventional techniques and a new nondifferential approach. In 10th European Signal Processing Conference, Vol. 2, Islamabad, Pakistan, pp. 921–924, 2000.

[4] Manjunath, B. S., Wu, P., Newsam, S., and Shin, H. D. A Texture Descriptor for Browsing and Similarity Retrieval, Journal of Signal Processing: Image Communication, Volume 16, Issue 1-2, page 33-43, 2000.

[5] Hejazi M. Shevlyakov R., G., and Ho Y.S., Modified discrete Radon transforms and their application to rotation-invariant image analysis, In Eighth IEEE Workshop on Multimedia Signal Processing (MMSP2006), Victoria, Canada, pp. 429–234, October 2006.

[6] Jafari-Khouzani K. and Soltanianzadeh H., Radon transform orientation estimation for rotation invariant texture analysis, IEEE Trans. Pattern Anal. and Machine Intell. 27, pp 1004–1008, 2005.

[7] Hejazi Mahmoud R., Ho Yo-Sung, An Efficient Approach to Texture-Based Image Retrieval. Int. J. Imaging Syst. Technol., Vol. 17, pp. 295–302, 2007.

[8] Asatryan D., Egiazarian K., Quality Assessment Measure Based on Image Structural Properties. Proc. of the International Workshop on Local and Non-Local Approximation in Image Processing, Finland, Helsinki, pp. 70-73, 2009.

[9] Cramer H., Mathematical Methods of Statistics, 1961.

[10] Egiazarian K., Saramäki T., Chugurian H., and Astola J., Modified B-spline filters and interpolators, in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP'96, Atlanta, Georgia, USA, vol.3, pp. 1743-1746, 1996.

Authors' Information

David Asatryan – *Professor, Head of group of the Institute for Informatics and Automation Problems of NAS Armenia, 1, P.Sevaki Str., 0014, Yerevan, Armenia, e-mail: dasat@ipia.sci.am.*

Major Fields of Scientific Research: Digital signal and image processing.

Karen Egiazarian – *Professor, Head of Transforms and Spectral Techniques Group, International Center for Signal Processing, Tampere University of Technology, P.O.Box 553, FIN-33101 Tampere, FINLAND, e-mail: karen.egiazarian@tut.fi.*

Major Fields of Scientific Research: applied mathematics, signal/image processing, and digital logic.

Vardan Kurkchiyan – *Ph.D student, Russian-Armenian (Slavonic) University, 123, Hovsep Emin Str., 0051, Yerevan, Armenia, e-mail: vakur@gmail.com.*

Major Fields of Scientific Research: Digital image processing.

PREDICTIVE MODELING OF SPATIAL REDISTRIBUTION IN DYNAMICAL MODELS OF GLOBAL VEGETATION PATTERNS UNDER CLIMATE CHANGE

Nelli Ajabyan

Abstract: *The paper discusses stability concepts and methods of stability investigation in models describing the climate-biosphere interaction for a hypothetical zero-dimensional model. The climate-biosphere model proposed by [Svirezhev-& von Bloh, 1997], exhibited interesting features concerning the stability and sensitivity to perturbations of the global system. The work is focused on the determination of transition times between the equilibriums due to random perturbations. In investigate bifurcations in the model to describe complex behavior in the model with multiple equilibriums. Global vegetation pattern dynamics strongly depends on production and hence evolutionary models must explicitly use this concept. We have shown analytically that GVP dynamics could demonstrate complex behavior with multiple equilibria, chaotic dynamics associated with them as well as transition zones between consequent states.*

Keywords*: global vegetation pattern, modeling*

ACM Classification Keywords J.2 PHYSICAL SCIENCES AND ENGINEERING, G.1.7 Ordinary Differential Equations: *Chaotic systems*

Introduction

Global vegetation pattern dynamics strongly depends on production and hence evolutionary models must explicitly use this concept. The description of global vegetation dynamics explicitly including the production as parameter was suggested in [Svirezhev Y. 2000]. In this model the influence of different factors on vegetation dynamics on concentrated in the function that describes the dependence of production on them. The evolution of GVP under climate change is described by the change of model parameters. It was shown that using the formalism of Lotka-Volterra model for competing species it revealed possible to construct some discrete structures remaining meanwhile in the framework of a continuous model description. In this work we will investigate bifurcations in the same model to describe complex behavior in the model with multiple equilibria.

Competition models, including diffusive Lotka-Volterra models, proved to be subject of intensive exploration recently. The neutral stability theory has been widely applied to modeling of species richness in tropical forests. One model developed by Tilman [Tilman, 1994] was used to give alternative explanation to that of neutral theory.

Bampfylde [Bampfylde, 2005] adapted Tilman's competition model to gain deeper understanding of mechanisms for coexistence of many species exhibited in rain forests.

Vandermeer [Vandermeer, 2004] provided examples demonstrating complicated patterns in tree species distribution of tropical rain forests. He proposed that transient dynamics represents the continual shuffling among various basins of attraction under perturbation caused by disturbance events repeatedly affecting the forest. The most interesting point here, as it is noted in [Bampfylde, 2005], is finding empirical evidence since the recruitment limitation (defined as the failure of species to colonize a suitable vacant site) is acute in forests when generation times are many decades.

An attempt is made for analytical investigation of special class of Lotka-Volterra competition models directed to description of global vegetation pattern proposed by Svirezhev [Svirezhev, 2000]. The state variables are interpreted as different types of vegetation which correspond to the density of living biomass of a selected type at some geographical point (x, y). We study the change of dynamics of model with respect to discrete values of parameters with further analysis of stability under continuous in some vicinity of a selected value. Under this condition the model can be described by ordinary differential equations. A movement in the transition zone between types of vegetation then is determined in parametric space rather than considering solutions of partial differential equation system. The most important is GVP map, which represents a geographical distribution of different types of vegetation. As it is indicated in [Svirezhev Y. 2000] the map is a space of discrete structure. To interpret the results we will use the approach developed for the original model, based on definition of a mapping from parameter space to geographical points.

The complex behavior is formally determined in the general dynamic model with three species when unperturbed system can itself possess chaotic orbits. The third component is included in the model as a pseudo-type designed to integrate different disturbances in the system, examples of which present seasonal growth of parasite insects or insect invasion, consequences of fires that cause a relatively long term impact on the change of dynamics. The resulting system hence is a dynamic system of one predator with two competing species as a resource type. In this model free spaces which occur after some events can be latter occupied by the species with faster response to occurrence of fruiting conditions. In a long term scale later it can be replaced by the second competitor, thus moving the border between species. The latter is not necessarily a straight line on a specific spot but could be a line of wavelike form.

The method of investigation is based on identifying invariant sets of the dynamical system. The smallest of them are the equilibria. The stability of an equilibrium can change in result of bifurcations as parameters of the system change. New invariant sets defined in result through a prescribed tuning of multiple parameters in the system. These sets are idealizations but are associated with a range of definite behavior in case of slightly weaker assumptions on parameters than the tuned ones. In more than two-dimensional case the invariant sets are destroyed by local or global bifurcations. Local bifurcations may generate periodic solutions or new homoclinic or heteroclinic cycles, while global destroy links between two invariant sets. It has been recognized recently (see, for

example, J. Porter, 2004, Magnitski 2004) that heteroclinic cycles are associated with bursting and intermittent behavior, as well as give rise to chaotic dynamics. In general the trajectories escape from the invariant sets when small symmetry-breaking terms are considered or under stochastic perturbations. A weakly broken symmetry gives rise to intermittent dynamics in some nearby invariant region, while noise may allow the trajectories to "jump across" the invariant space. Due this process of potential random switching the partition of phase space and definition of transitions through it is very important for predicting the behavior of the system

Routes to chaos and parameter combinations in which there exist alternative limit cycles need detailed examination to reveal whether the jumps in dynamics will be attained in a real system.

1. Model description

To consider dynamics of two-species model we assume, following approach in [Svirezhev Y. 2000], that types of vegetation are ordered along a special coordinate x, for instance along a meridian. Then the model is described by the following system:

$$\frac{\partial B_1}{\partial t} = B_1(\varepsilon_1(x) - \gamma_{11}B_1 - \gamma_{12}B_2)$$
$$\frac{\partial B_2}{\partial t} = B_2(\varepsilon_2(x) - \gamma_{21}B_1 - \gamma_{22}B_2) \qquad (1.1)$$

where $\varepsilon_i(x) = \varepsilon_i(T(x), H(x)) = g_i(T(x), H(x)) - m_i = g_i(x) - m_i$

We will further adapt the model by assuming ε_i being a parameter that can vary in some limits, but we exclude direct indication of dependence from spatial coordinate *x*. Then the species compete for a site the characteristics of which are specified by ε_i . In this system one type of vegetation could replace the other or they could coexist, the dynamics of whole area is determined by the dynamics of spots, the system is similar to that of the metapopulation model. The next change consists in introduction of scaled variables defined by the following:

$$\gamma_{11}B_1 \to \tilde{B}_1 \qquad a_2 = \frac{\gamma_{12}}{\gamma_{11}} \qquad \gamma_{11} = \gamma_{22} = 1$$
$$\gamma_{22}B_2 \to \tilde{B}_2 \qquad a_1 = \frac{\gamma_{21}}{\gamma_{11}} \qquad \gamma_{12} = a_2,\ \gamma_{21} = a_1$$

However we will use the same notation B_i for the sake of simplicity, then the model is described by the following autonomous system of ordinary differential equations

$$\frac{dB_1}{dt} = B_1(\varepsilon_1 - B_1 - a_2B_2)$$
$$\frac{dB_2}{dt} = B_2(\varepsilon_2 - a_1B_1 - B_2) \qquad (1.2)$$

The state (0, 0) is a steady state for (1.2). i.e. dead spaces without any vegetation or free gaps are assumed in this model. The non-trivial equilibria of the system are:

1. $\{B_1^{(1)},0\}$ $B_1^{(1)} = \varepsilon_1$
2. $\{B_2^{(1)},0\}$ $B_2^{(1)} = \varepsilon_2$
3. $\{B_1^{(2)}, B_2^{(2)}\}$ $B_1^{(2)} = \frac{\varepsilon_2 - a_2\varepsilon_1}{1-a_1a_2}$ $B_2^{(2)} = \frac{\varepsilon_1 - a_1\varepsilon_2}{1-a_1a_2}$

From the linear stability conditions the first equilibrium is a stable node if both ε_i are negative since the eigenvalues of the system at this equilibrium are $\lambda_i=\varepsilon_i$ and it is unstable if ε_i are positive. In fact positive values provide growth for both competitors, which is limited by carrying capacity of a spot or growth thresholds for the competing species.

The origin is dead space if ε_i are zero; when they are negative the system to asymptotically stable but dead biomass state. The eigenvalues calculated for the second fixed point $\{B_2^{(1)},0\}$ are $\lambda_1 = \varepsilon_1 - a_2\varepsilon_2,\ \lambda_1 = -\varepsilon_2$

The table summarizes the description of stability regions for the state $\{B_2^{(1)},0\}$

$\varepsilon_2 > 0$ $\varepsilon_1 > 0$	$\varepsilon_1 < a_2\varepsilon_2$	asymptotically stable coexistence	
	$\varepsilon_1 > a_2\varepsilon_2$	saddle point unstable B_1 is dominant	
	$\varepsilon_1 = a_2\varepsilon_2$	the stability is determined by nonlinear part	
$\varepsilon_2 < 0$ $\varepsilon_1 > 0$	$\lambda_2 > 0$ $\lambda_1 > 0$	unstable	species coexistence
$\varepsilon_2 < 0$ $\varepsilon_1 < 0$	$\lambda_2 > 0$ $\lambda_1 = \varepsilon_1 + a_2\varepsilon_2$ $\lambda_1 < 0,\ a_2 \mid \varepsilon_2 \mid < \varepsilon_1$	saddle point unstable B_1 dominant	
	$a_2 \mid \varepsilon_2 \mid > \varepsilon_1$	$\lambda_1 > 0, \lambda_2 > 0$ both species develop	
$\varepsilon_2 > 0$ $\varepsilon_1 < 0$	$\lambda_2 < 0,\ \lambda_1 > 0$	asymptotically stable $\lvert\varepsilon_1\rvert < a_2\varepsilon_2$	
	$a_2 \mid \varepsilon_2 \mid < \varepsilon_1$	unstable	

By symmetry we can simply describe the state $\{B_1^{(1)}, 0\}$ eigenvalues for which are $\lambda_1 = -\varepsilon_1$, $\lambda_2 = \varepsilon_2 - a_1\varepsilon_1$ asymptotically stable

1. $\varepsilon_i > 0, \varepsilon_2 < a_1\varepsilon_1$

2. $\varepsilon_i < 0, \varepsilon_2 < 0, \lambda_i > 0$ Unstable

3. $\varepsilon_i < 0$ saddle point if $\varepsilon_2 > a_1 \mid \varepsilon_1 \mid$ otherwise $\lambda_i > 0$

The characteristic equation for ($B_1^{(2)}, B_2^{(2)}$) is

$$\lambda^2 + \lambda \frac{\varepsilon_2(a_2 - 1) - \varepsilon_1(a_1 - 1)}{a_1 a_2 - 1} + M = 0$$

Where $M = \varepsilon_1\varepsilon_2 + \frac{a_1\varepsilon_1^2}{a_2a_1 - 1} + \frac{a_2\varepsilon_2^2}{a_2a_1 - 1} + \frac{\varepsilon_1(a_2a_1 + 1)}{a_2a_1 - 1}$

Then the bifurcation parameter is: $M_c = 0$ if $a_2 = a_1 \neq 1$

The parameter can become zero under

$$\varepsilon_2 = -\varepsilon_1 \;\&\; a_2 = a_1, M_c = \varepsilon_1 + \varepsilon_2 - (a_1\varepsilon_1 + a_2\varepsilon_2), \lambda_i = -\frac{1}{2}M_c \pm \sqrt{\frac{M_c^2}{4} - M}$$

The system will have purely imaginary eigenvalues $\pm i\sqrt{M}$ at μ=0, since the values $a_1 = a_2 = 1$ are restricted and interspecific competition parameters are always positive M will be always positive, but it can become zero only at ε_I =0 for both *i*.

Thus when ε_i take values on the straight line l the limit cycle arises from Hopf bifurcation (Figure 1.).

We can check the condition $\frac{\partial \operatorname{Re}\lambda(\mu)}{\partial\mu} > 0$, where $\frac{\partial \operatorname{Re}\lambda(\mu)}{\partial\mu} = \varepsilon_1 + \varepsilon_2 - (a_1\varepsilon_1 + a_2\varepsilon_2)$.

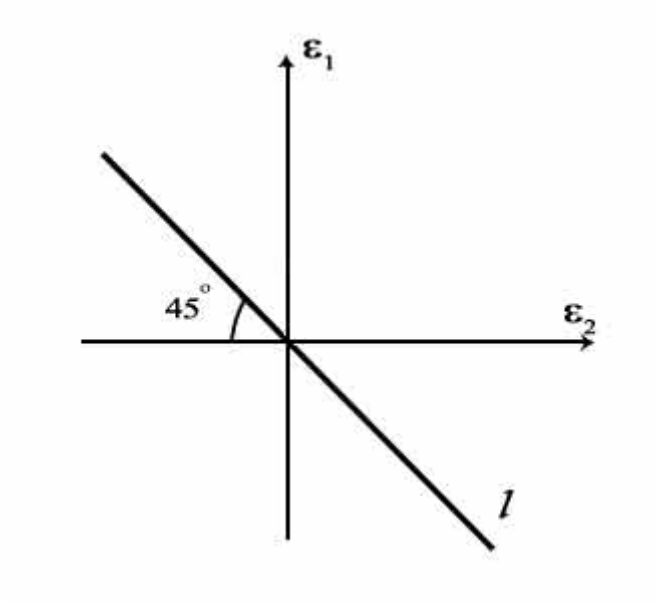


Figure 1.

For the points in part I the expression will be positive, then the system has a stable focus, while in the upper part II it will become negative, hence $\lambda_i > 0$. This means that at μ=0oscillatory solutions occur in the system with an amplitude proportional to $\sqrt{\rho}$ and the period $T \approx \frac{2\pi}{\vartheta}$, where $\lambda_i = \rho \pm i\vartheta$. At ε_i =0 this point becomes a point of center type and it is attracting from one side and repelling from the other. In the vicinity of the limit cycle the system can be brought to the normal form defined as

$$\dot{B}_1 = -\vartheta B_2 + [\rho \pm (B_1^2 + B_2^2)]B_1$$
$$\dot{B}_2 = -\vartheta B_1 + [\rho \pm (B_1^2 + B_2^2)]B_2$$

where the signs + ,– stand for subcritical and supercritical bifurcations accordingly. By estimating the nonlinear component for B_i it is possible to prove that the bifurcation is supercritical in large interval of parameter values. The corresponding transition is soft since a continuous change occurs.

Consider the transition in the system for the particular case with parameter values determined by

$$\gamma_{12} = a_2 = 2,\ \gamma_{21} = a_1 = 2$$

The system in nontrivial equilibrium $(B_1^{(2)}, B_2^{(2)})$, $B_1^{(2)} = \frac{2\varepsilon_2 - \varepsilon_1}{3}, B_2^{(2)} = 2\varepsilon_1 - \varepsilon_2$ has Jacobi matrix

$$\begin{vmatrix} \frac{-7\varepsilon_1 + 4\varepsilon_2}{3} & \frac{\varepsilon_1 - 2\varepsilon_2}{3} \\ -2(2\varepsilon_1 - \varepsilon_2) & \frac{8\varepsilon_2 - 10\varepsilon_1}{3} \end{vmatrix}$$

$$\frac{dB_1}{dt} = B_1(\varepsilon_1 - B_1 - 2B_2)$$
$$\frac{dB_2}{dt} = B_2(\varepsilon_2 - 2B_1 - B_2)$$

$$\lambda^2 - \lambda\frac{12\varepsilon_2 - 17\varepsilon_1}{3} + \frac{82\varepsilon_1^2 + 44\varepsilon_2^2 - 106\varepsilon_1\varepsilon_2}{9} = 0$$

The eigenvalues are

$$\lambda_{1,2} = -\frac{17\varepsilon_1 - 12\varepsilon_2}{6} \pm \sqrt{\left(\frac{17\varepsilon_1 - 12\varepsilon_2}{6}\right)^2 - \frac{82\varepsilon_1^2 + 44\varepsilon_2^2 - 106\varepsilon_1\varepsilon_2}{9}} \qquad (1.3)$$

From (3) it is clear that bifurcation of the limit cycle is possible if $17\varepsilon_1 = 12\varepsilon_2$

At this point the has a pair of imaginary eigenvalues given by

$\lambda_i \approx \pm\varepsilon_1 i\sqrt{21{,}6}$, where $\varepsilon_2 = \frac{17}{12}\varepsilon_1 \approx 1{,}4\varepsilon_1$

The necessary conditions for Hopf bifurcation hold, the imaginary part will not turn to zero unless $\varepsilon_2 \neq 0$.

2. Routes to chaos in two-species dynamic model with disturbance factors.

To consider dynamics of two-species model we assume, following the approach in [Svirezhev, 2000], that types of vegetation are ordered along a special coordinate x, for instance along a meridian. We will further adapt the model by assuming ε_i being a parameter that can vary in some limits, but we will exclude direct indication of dependence from spatial coordinate *x*. Then the species compete for a site the characteristics of which are specified by ε_i. In this system one type of vegetation could replace the other or they could coexist, the dynamics

of whole area is determined by the dynamics of spots, the system is similar to that of the metapopulation model. The next change consists in introduction of scaled variables defined by the following:

$$\gamma_{11}B_1 \to \tilde{B}_1 \qquad a_1 = \frac{\gamma_{21}}{\gamma_{11}} \qquad \gamma_{11} = \gamma_{22} = 1$$
$$\gamma_{22}B_2 \to \tilde{B}_2 \qquad a_1 = \frac{\gamma_{12}}{\gamma_{11}} \qquad a_2 = \gamma_{12},\ \gamma_{21} = a_1$$

However we will use the same notation B_i for the sake of simplicity, then the model is described by the following autonomous system of ordinary differential equations:

$$\begin{aligned} \dot{B}_1 &= B_1(\varepsilon_1 - B_1 - a_2B_2) \\ \dot{B}_2 &= B_2(\varepsilon_2 - B_2 - a_1B_1) \end{aligned} \qquad (2.1).$$

To define shifts in dynamics of the two-species model generated by factors that can make severe impact such as fires we introduce a new variable B_3. We assume B_3 time dependent, dynamics of which is given by the equation

$$B_3(t) \cong e^{\int_t f_3(B_1,B_2)dt}$$

We assume B_3 of the order ε but are interested in impact on the system dynamics provided on relatively short time intervals. Since f_3 is a vanishing function of order ε we will apply averaging and write

$$B_3(t) \cong e^{\beta t}\{const + \frac{1}{T}\int_0^t f_3(B_1, B_2)dt\}$$

The three variable systems now can be written in the form

$$\begin{aligned} \dot{B}_1 &= B_1(\varepsilon_1 - B_1 - a_2B_2 - \gamma_{13}B_3) \\ \dot{B}_2 &= B_2(\varepsilon_2 - B_2 - a_1B_1 - \gamma_{23}B_3) \\ \dot{B}_3 &= B_3(\beta - \gamma_{31}B_1 - \gamma_{32}B_2) \end{aligned} \qquad (2.2).$$

The parameters γ_{i3} characterize the strength of B_3 impact on species B_i, γ_{3i} the species survival potential.

The state (0,0,0) is a steady state for (2.2). i. e. dead spaces without any vegetation or free gaps are assumed in this model. The parameter β can be positive and then has interpretation in terms of intensity of the impact, while

negative values correspond to the damping rate of the disturbance. We will show that one of the equlibria of the system is a saddle-focus, while two others can be stable, unstable nodes or saddle-nodes.

The three equilibria of the system along $B_3 = 0$ are:

$$1.\{B_1^{(1)}, 0, 0\}, \quad 2.\{0, B_2^{(1)}, 0\}, \quad 3.\{B_1^{(2)}, B_2^{(2)}, 0\}$$

Where $B_1^{(1)} = \varepsilon_1, \quad B_2^{(1)} = \varepsilon_1, \quad B_2^{(1)} = \dfrac{\varepsilon_2 - a_1\varepsilon_1}{1 - a_1 a_2}, \quad B_1^{(2)} = \dfrac{\varepsilon_1 - a_2\varepsilon_2}{1 - a_1 a_2}$

The eigenvalues for the first equilibrium are $\lambda_3 = \beta - \gamma_3\varepsilon_1, \quad \lambda_1 = \varepsilon_1, \ \lambda_2 = \varepsilon_1 - a_2\varepsilon_2$.

From these expressions we can see that the eigenvalues are always real, but the signs are determined from relations among B_1, ε_1, ε_2.

The characteristic equation of the system for a non-trivial equilibrium can be easily defined from the linear stability matrix evaluated at {B_1, B_2, 0}.

$$\begin{vmatrix} \dfrac{a_2\varepsilon_2 - \varepsilon_1}{1 - a_1 a_2} & a_2\dfrac{a_2\varepsilon_2 - \varepsilon_1}{1 - a_1 a_2} & -\gamma_{13}\dfrac{\varepsilon_1 - a_2\varepsilon_2}{1 - a_1 a_2} \\ a_1\dfrac{a_1\varepsilon_1 - \varepsilon_2}{1 - a_1 a_2} & \dfrac{a_1\varepsilon_1 - \varepsilon_2}{1 - a_1 a_2} & -\gamma_{23}\dfrac{\varepsilon_2 - a_1\varepsilon_1}{1 - a_1 a_2} \\ 0 & 0 & \beta - \gamma_{31}\dfrac{\varepsilon_1 - a_2\varepsilon_2}{1 - a_1 a_2} - \gamma_{32}\dfrac{\varepsilon_2 - a_1\varepsilon_1}{1 - a_1 a_2} \end{vmatrix} \qquad (2.3)$$

It is evident that the system has one real eigenvalue λ_3 and two complex-conjugate λ_{12}

$$\lambda_3 = \beta - \frac{\gamma_{31}(\varepsilon_1 - a_2\varepsilon_2) + \gamma_{31}(\varepsilon_2 - a_1\varepsilon_1)}{1 - a_1 a_2} \qquad \lambda_{1,2} = \rho \pm i\vartheta$$

Where $$\vartheta = \frac{1}{2(1 - a_1 a_2)}\sqrt{M}, \quad \rho = -\frac{1}{2}\left[\frac{\varepsilon_1(1 - a_1) + \varepsilon_2(1 - a_2)}{1 - a_1 a_2}\right] \qquad (2.4)$$

$$M = 4(1 - a_1 a_2)\left[\varepsilon_1\varepsilon_2(1 + a_1 a_2) - a_1\varepsilon_1^2 - a_2\varepsilon_2^2\right] - (\varepsilon_1 - \varepsilon_2 + a_1\varepsilon_1 - a_2\varepsilon_2)^2,$$

We assume $a_1 = 1 \,\&\, a_2 = 1, \ M > 0.$

The emergence of a two-dimensional stable manifold is a consequence of the stable limit cycle existence in the two-dimensional system with parameter values $\varepsilon_1 = \frac{1-a_2}{a_2-1}\varepsilon_2$. This relation means that ε_1 and ε_2 must have different sings providing a_i are positive. Let's remind that $\varepsilon_i(x)$ is $\varepsilon_i(x) = g_i(x) - m_i$ present the difference between the annual productivity and mortality. The shift of border means that destructive process prevails over reproductive. However, by translating the origin into point $(B_1^{(2)}, B_2^{(2)}, 0)$ more transparent conditions for variables in the transition zone. Let B_1^*, B_2^* be

$$B_1^* = B_1 - B_1^{(2)}$$
$$B_2^* = B_2 - B_2^{(2)}$$

Clearly, new B_i^* could be either positive or negative, but we can include as the direction of shift. Suppose the shift from the forest to steppe or to bare soil temporarily, it is possible to write explicitly

$$B_1^* = B_1 - B_1^{(2)}$$
$$B_2^* = B_2 + B_2^{(2)}$$

Then the Jacobian matrix (2) will take the form

$$\begin{vmatrix} \varepsilon_1 & -a_1 B_1^{(2)} & -\gamma_{13}\dfrac{\varepsilon_1 - a_2\varepsilon_2}{1-a_1a_2} \\ a_1 B_2^* & \varepsilon_2 & -\gamma_{23}\dfrac{\varepsilon_2 - a_1\varepsilon_1}{1-a_1a_2} \\ 0 & 0 & \beta - \gamma_{31}\dfrac{\varepsilon_1 - a_2\varepsilon_2}{1-a_1a_2} - \gamma_{32}\dfrac{\varepsilon_2 - a_1\varepsilon_1}{1-a_1a_2} \end{vmatrix}$$

Shil'nikov's theorem states the existence of Smale's horseshoe and related chaotic irregular dynamics in the vicinity of the robust saddle- focus, $\lambda_3 > 0 \,\&\, \rho < 0$. As it is emphasized in [[Magnitski, 2004, p. 92] this theorem does not prove the existence of an invariant set specific the horseshoe mapping. Following the elaboration in [Magnitski 2004] we bring the model equations to the form

$$\begin{aligned} \dot{B}_1 &= \rho B_1 - \vartheta B_2 + P(B_1, B_2, B_3) \\ \dot{B}_2 &= \vartheta B_1 + \rho B_2 + Q(B_1, B_2, B_3) \\ \dot{B}_3 &= \lambda_3 B_3 + R(B_1, B_2) \end{aligned} \qquad (2.5)$$

where P, Q, R some functions containing the components from the second power. The system has two manifolds: a stable and unstable ones. The chaotic dynamics emerges in case $|\rho| < \lambda_3$ unless the speed of movement away from the fixed point on unstable manifold is greater than the speed of attraction to the point.

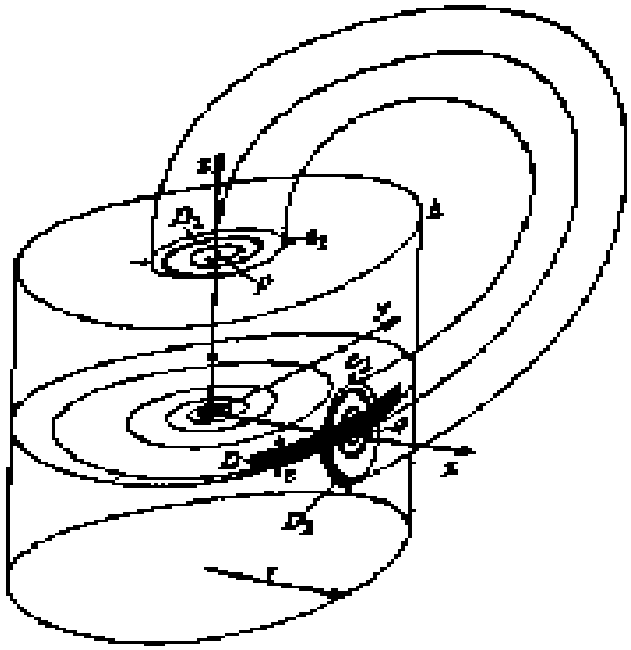


Fig. 2 illustrates the change of trajectories in the phase space (adopted from [Magnitski, 2004]).

Let D be some region on the surface of the cylinder, the points in which are specified with two variables ξ, θ, where $0 < \xi < \theta, \ |\theta| < \theta_{\max}$ is the angle between the old position of some point q and the transferred one. Since the linear approximation is valid inside the cylinder the solutions of the three-dimensional system are defined by

$$\dot{B}_1 = r\exp(\rho t)\sin(\vartheta t + \theta)$$

$$\dot{B}_1 = r\exp(\rho t)\cos(\vartheta t + \theta)$$

$$\dot{B}_1 = \xi\exp(\rho t)\sin(\lambda_3 t)$$

On the starting point the movement along the separatrix will bring some point q from D to a new area D_1 situated on the upper surface of the cylinder. The size of the region is restricted by a cycle of a diameter $\varepsilon_1 = 2r\left(\frac{\varepsilon}{h}\right)^{\frac{|\rho|}{\lambda_3}}$. On the following transition the central point p of D_1 makes transition to the original position q.

The mapping of the whole D_1 to new D_2 can be approximated by a spiral lying inside the cycle with the diameter of *k* fold the one of D_1, i.e. $\varepsilon_2 = 2\pi r\left(\frac{\varepsilon}{h}\right)^{\frac{|\rho|}{\lambda_3}}$

When $\varepsilon_1 < 2\varepsilon$ and small ε the spiriling curve will remain in D while at $\frac{|\rho|}{\lambda_3} < 1$ change $\varepsilon_1 > 2\varepsilon$ specific for the horseshoe mapping occurs.

In the limit $\varepsilon \to 0$ a countable number of horseshoes appear, but introduction of small perturbations will leave a finite number of horseshoes persistent.

Let us select two strips V_1, V_2 in the population area each containing the points that belonged to the state at initial moment, this could be biomasses with values in some attracting zone of the limit cycle then the second picture we can present the set of points survived after two steps of the iteration process. There can be a sequence of iterations from the formal point of view. In the model the iteration correspond to some shifts that then could occur in reverse direction. Under decay or some slow recovery from the disturbance can the species can return in time run to some state close to the initial or with a distorted shape of the border.

Another version of Shil'nikov's theorem is valid for the case $\rho > 0, \lambda_3 < 0$, where the condition for the emergence of chaotic dynamics is $\rho < |\lambda_3|$

The point now is to consider the behavior of the system with $\rho = \lambda_3$. The case with both ρ, λ_3 positive corresponds to the emergence of a unstable focus, in fact this means that the third component is not disturbing, it could be even fruiting. It is not feasible for GVP model in this approach and we will not take it into account. Thus we shall consider the conditions under which the trajectories of the system will "jump" across the invariant subspace of saddle focus stuck to another state, thus bringing the original model to destruction of the existing pattern and replacing it by new one. The boundary which has just been the intersection of the stable and unstable manifolds in the consequence of perturbations now will become a connection between two fixed points. To define the properties of the heteroclinic cycle separatrix that emerges in result of this new non local bifurcation. Let us consider the change in parameter values that can push the system to a new bifurcation.

The transition along the connecting manifold can correspond either to substantial growth for one of the species or the number and size of free spots growth. The opportunities for further "recolonization" of the spot depend however from the selected species potential to react on new conditions and not only the absence of the competitor, as well as disturbing factor.

From the expressions for $B_1^{(2)}, B_2^{(2)}$ it is easy to conclude that the system can potentially come to $(0, B_2^{(1)}, 0)$ when $\varepsilon_1 \to a_2\varepsilon_2$ or to the state $(B_1^{(1)}, 0, 0)$ in case $\varepsilon_2 \to a_1\varepsilon_1$.

Another point is made by a change in the sign of M, after which the system will have real eigeinvalues. Consider first the type of the equilibrium $\{0, B_2^{(1)}, 0\}$.

Let's put $\delta = \varepsilon_1 - a_2\varepsilon_2$ then the eigenvalues are:

$$\lambda_1 = \delta, \lambda_2 = -\varepsilon_2, \lambda_3 = \beta - \gamma_{32}\varepsilon_2$$

Assuming $\varepsilon_2 > 0,\ \beta > 0, \beta > \gamma_{32}\varepsilon_2$ the state will be a saddle node with two-dimensional stable node and one-dimensional unstable.

There are two points in the system: a saddle-focus with one-dimensional unstable and two-dimensional stable and a saddle-node with the contrary. Based on numerical investigations it can be concluded [Magnitski 2004, p. 93] that the contour connecting these two points is not robust. An irregular dynamics is observed in the neighborhood of the bifurcation point with infinitely many stable and unstable limit cycles. A special closed contour connecting the saddle focus with saddle –node emerges at $\varepsilon_i \to 0$. From the model point of view this means that the presence of a disturbing factor will continue to provide impact bringing to oscillations in the system.

To consider global bifurcations let's introduce two parameters: δ for the difference between ρ and $\lambda_{3,}$ d_{12} and d_{21} for the coefficients at in the nonlinear part of the normal form and write the equations in the form:

$$\dot{B}_1 = -\vartheta B_2 + \rho B_1 + d_{12}(B_1^2 + B_2^2)B_1$$

$$\dot{B}_2 = -\vartheta B_1 + \rho B_2 + d_{21}(B_1^2 + B_2^2)B_2,\ \ \delta = \lambda_3 - \rho,\ \ B_1^2 + B_2^2 = r^2$$

$$\dot{B}_3 = (\delta + \rho)B_3$$

When $\delta = 0$ Hopf bifurcation is supercritical unless $d_{1,2}$ and $d_{2,1}$ are negative, when the signs of d_{ij} change (under the impact of disturbance) the bifurcation will become subcritical. The bifurcation diagram (Fig. 3) shows the scheme for the upper part of parameter space partition, with λ_3 >0.

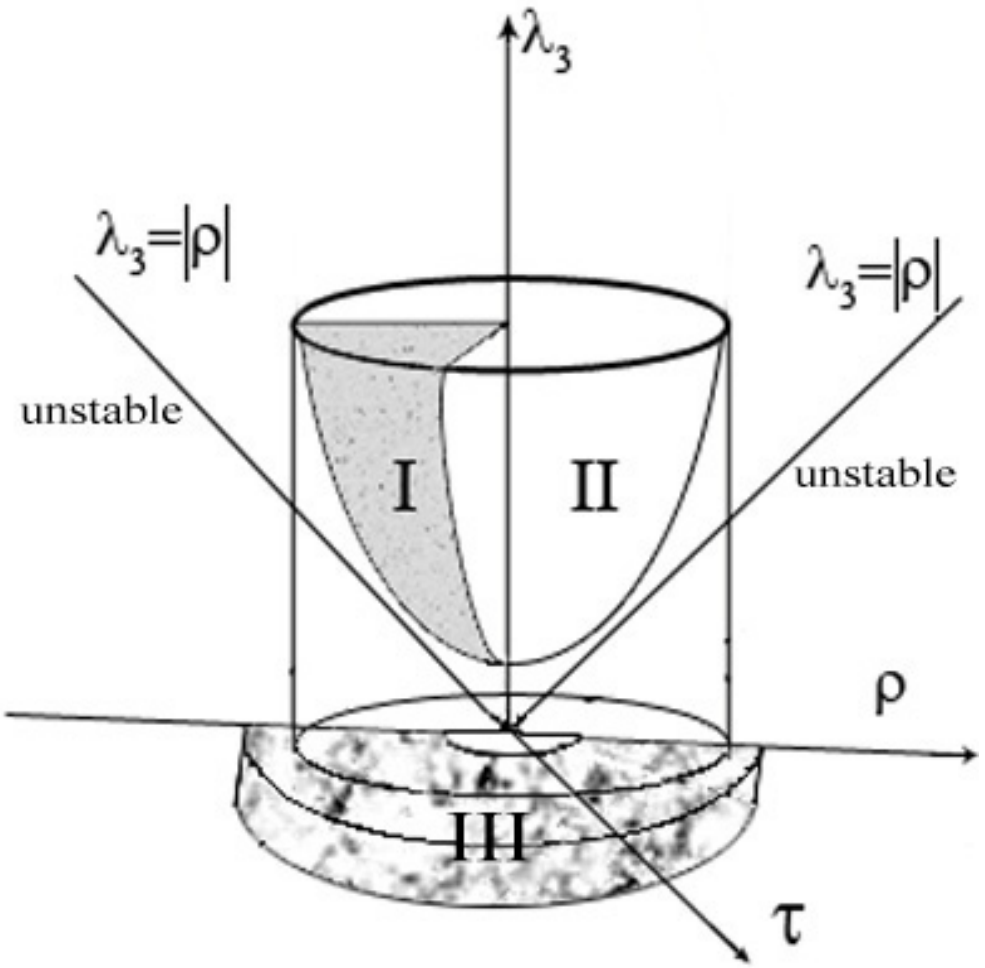


Fig.3. The part of the conic surface shows the region where Silnikov's type chaotic dynamics emerges. The bisectors divide the quadrants to the parts based on restrictions $\lambda_3 = -\rho,\ \lambda_3 = \rho$ .Region III is marked with Hopf bifurcation.

The chaotic dynamics is associated with some type of local or global bifurcation. When δ is small and λ_3 is negative the limit cycle created by Hopf bifurcation surrounding the nontrivial equilibrium on the phase portrait of the system persists. With a change in λ_3 and, hence the increasing disturbance, a saddle-focus bifurcations take place. At the border (line l_1) in some small part defined. While the region enlarges a new closed loop encircling both saddle focus and saddle node equilibrium emerges. At the saddle node one of the eigenvalues of the system is negative and the second is positive. After the saddle-node gets destroyed, the new emergent dynamics with $\{B_1^{(1)}, 0\}$ makes the equilibrium relatively stable, but with only one of the species survived. When the destabilizing impact of B_3 decreases or even disappears, the new distribution will depend on several factors.

Thus, if a small collection of the removed species seedlings remains and the survivors can use new beneficial conditions (such as free gaps, for example) the system has an opportunity to restore the primary border.

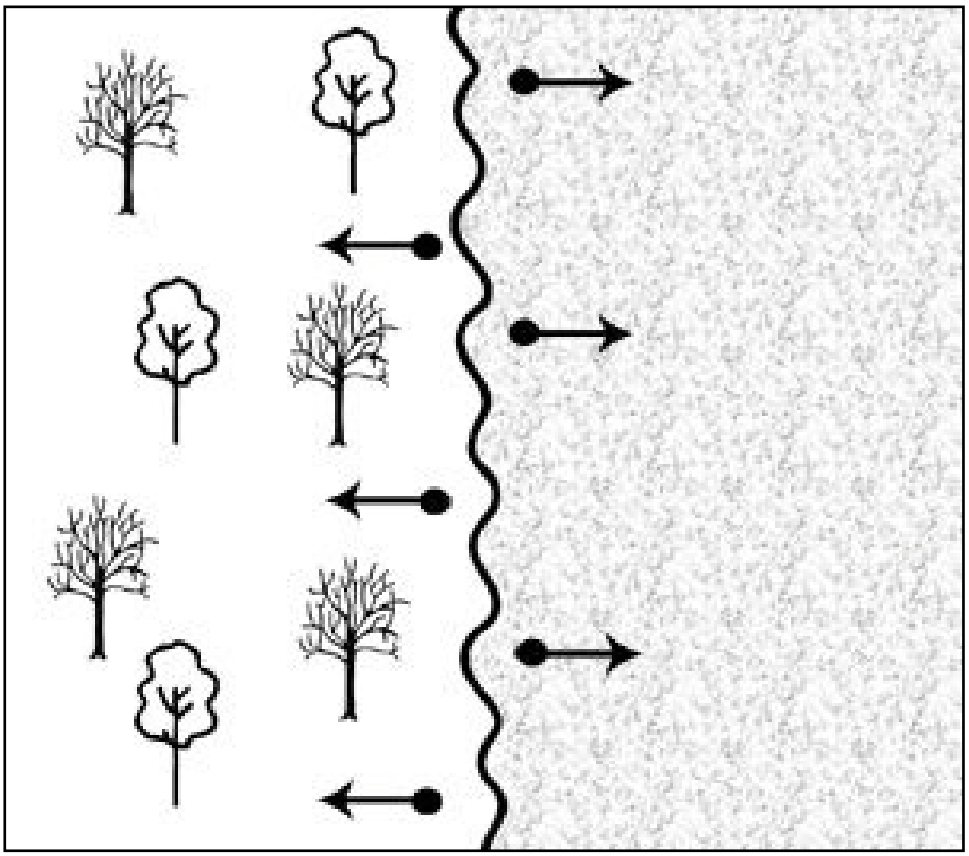


Fig.4.The line shows the border between the part covered with the woods and the steppe.

Conclusion

We have shown analytically that GVP dynamics could demonstrate complex behavior with multiple equilibria, chaotic dynamics associated with them as well as transition zones between two consequent states. The main parameter of the model is defined by distinct values of productivity function. A new variable has been integrated to the system; only general assumptions are made with respect to this function characterization. We model the competition between the destabilizing effects caused by some event with stabilizing potential of the two-species system. We illustrate that in case the strength of destabilization is weak (with respect to existing resource, i.e. biomasses and their potential to renew and survive) and then damping takes place the chaotic dynamics is revealed that can provide a distortion of the borderline or an essential shift, which however will not bring the system to the origin with corresponding dead space in the model.

Acknowledgements

The work was partially supported by DAAD grant Gr 6205-A0538516 and was implemented when the author was in PIC, Germany. The author expresses her gratitude for hospitality and helpful discussions.

Bibliography

[Svirezhev Y., 2000]. Svirezhev Y., 2000, Lotka-Volterra models and the global vegetation pattern. Ecological Modeling 135, pp.135-146.

[Svirezhev-von Bloh 1997]. Svirezhev Y. M., W. von Bloh. Climate, vegetation and global carbon cycle: the simplest zero-dimensional model. Ecological Modeling, 101, 1997 pp. 75-95.

[Bampfylde C. et al, 2005]. Bampfylde C., Brown N. D., Gavaghan D. J., Maini P. K., 2005, Modeling rain forest diversity: The role of competition. Ecological modeling, 188, pp. 253-278.

[Vandermeer J., 2004]. Vandermeer J., Granzow I., Perfecto I., Boucher D., Ruiz J., Kaufman A. 2004. Multiple basins of attraction in a tropical forest: evidence for nonequilibrium community structure. Ecology 85(2), pp. 575-579.

[Magnitski N.,2004].Magnitski N., Sidorov V., 2004. New methods of chaotic dynamics. Moscow (in Russian).

[Porter J., 2004].Porter J., 2004, Dynamics near robust heteroclinic cycles. In: Dynamics and bifurcations of patterns in dissipative systems. pp. 329-356. World Scientific Publishing Co.

[Tilman, 1994]. David Tilman, Competition and Biodiversity in Spatially Structured Habitats.. Ecology, Volume 75, Issue 1 (Jan., 1994), 2-16.

Authors' Information

Ajabyan Nelli –Researcher, Institute for Informatics and Automation Problems of NAS RA, e-mail: najabyan@ipia.sci.am.

Major fields of Scientific Research: predictive modeling of ecosystems, problems of assessment of response on disturbances.

A METHOD OF CONSTRUCTING PERMUTATION POLYNOMIALS OVER FINITE FIELDS

Melsik Kyureghyan, Sergey Abrahamyan

Abstract: *In this paper we consider the problem of characterizing permutation polynomials of the shape* $P(x) = x + \gamma f(x) + \delta\, g(x) + \tau l(x)$ *over the field* F_q*; that is, we seek conditions on the coefficients of a polynomial which are necessary for it to represent a permutation.*

Keywords: *finite field, permutation polynomial, linear translator.*

Introduction

Let q be a power of a prime number and F_{q^n} be the finite field of order $q^n \geq 1$. Recall that any mapping of a finite field into itself is given by polynomial. A polynomial $F(x)$ is called a permutation polynomial of F_{q^n} if it induces a permutation on F_{q^n}. These polynomials were first explored in the research of Betti [Betti,1851], Mathieu and Hermite [Hermite 1863] as a way of representing permutations. A general theory was developed by Hermite [Hermite 1863] and Dickson [Dickson 1896], with many subsequent developments by Carlitz et.al. The construction of permutation polynomials over any finite fields is a challenging mathematical problem. Interest in permutation polynomials stems from both mathematical theory as well as practical applications such as cryptography. Recent papers [Betti,1851]- [Markos 2011] highlight a method of construction of permutation polynomials. The given article considers permutations of the form $x + \gamma f(x) + \delta\, g(x) + \tau l(x)$ over F_q .

Preliminaries

Let's start with recalling some definitions and basic results that will be helpful to derive our main result.

Definition 1 Let $f: F_{p^n} \to F_p$ and $c \in F_p$. We say that $\alpha \in F_{p^n}^*$ is a c linear structure of the function f if $f(x+\alpha) - f(x) = c$ for all $x \in F_{p^n}$.

Note that if α is a c-linear structure of f, then necessarily $c = f(\alpha) - f(0)$.

Definition 2 Define $F(x) = G(x)^{\circ}H(x)$ composition of the mapping G with H.

Proposition1 ([Kyureghyan G. 2011**] Proposition1)** Let $\alpha, \beta \in \mathbb{F}_{q^n}^*$, $\alpha + \beta \neq 0$ and $a, b, c \in \mathbb{F}_q$, $c \neq 0$. If α is an a-linear translator and β is a b-linear translator of a mapping $f: \mathbb{F}_{q^n} \to \mathbb{F}_q$, then $\alpha + \beta$ is an $(a+b)$-linear translator of f and $c \cdot \alpha$ is a $(c \cdot a)$-linear translator of f. In particular, if $\Lambda^*(f)$ denotes the set of all linear translators of f, then $\Lambda(f) = \Lambda^*(f) \cup \{0\}$ is an $\mathbb{F}_q$-linear subspace of $\mathbb{F}_{q^n}$.

Proposition 2 ([Kyureghyan G. 2011**] theorem3)** Let $\gamma \in \mathbb{F}_{q^n}$ be a b-linear translator of $f: \mathbb{F}_{q^n} \to \mathbb{F}q$ and $b \neq -1$ then the inverce maping of the permutation $Fx = x + \gamma f x$ is

$$\mathrm{F}^{-1}(\mathrm{x}) = \mathrm{x} - \frac{\gamma}{\mathrm{b}+1}\mathrm{f}(\mathrm{x}).$$

Proposition 3 ([Kyureghyan G. 2011**] theorem8)** Let $\gamma \in \mathbb{F}_{q^n}$ be a b-linear translator of $f: \mathbb{F}_{q^n} \to \mathbb{F}q$.

(a) Then $F(x) = x + \gamma f(x)$ is a permutation of $\mathbb{F}_{q^n}$ if $b \neq -1$.
(b) Then $F(x) = x + \gamma f(x)$ is a q-to1 mapping of $\mathbb{F}_{q^n}$ if $b = -1$.

Proposition 4 ([Kyureghyan G. 2011**] theorem10)** Let $\gamma, \delta \in \mathbb{F}_{q^n}$. Suppose γ is a b_1-linear translator of $f: \mathbb{F}_{q^n} \to \mathbb{F}_q$ and a b_2-linear translator of $g: \mathbb{F}_{q^n} \to \mathbb{F}_q$, and moreover δ is a d_1-linear translator of f and a d_2-linear translator of g. Then

$$F(x) = x + \gamma f(x) + \delta g(x)$$

is a permutation of $\mathbb{F}_{q^n}$, if $b_1 \neq -1$ and $d_2 - \frac{d_1 b_2}{b_1+1} \neq -1$, or by symmetry, if $b_2 \neq -1$ and $d_1 - \frac{d_2 b_1}{b_2+1} \neq -1$.

Constructing Permutation

In this section we characterize permutation polynomials of the form

$$P(x) = x + \gamma f(x) + \delta\, g(x) + \tau l(x)$$

Theorem1

Let $F(x) = x + \gamma f(x) + \delta g(x)$ be a permutation polynomial in $\mathbb{F}_{q^n}$. Suppose γ is a b_1 –linear translator of $f: F_{q^n} \to F_q$ and a b_2 –linear translator of $g: F_{q^n} \to F_q$ moreover δ is a d_1 –linear translator of f and a d_2 –linear translator of g.

Then the inverse mapping of the permutation $F(x) = x + \gamma f(x) + \delta g(x)$

is $F^{-1}(x) = x - \left(f(x) - d_1 \frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}\delta$

where $A = (1 + d_2)(b_1 + 1) - d_1 b_2$.

Proof

Consider

$$F(x)^{\circ}\left(x - \left(f(x) - d_1 \frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}\delta\right)$$

$$= x - \left(f(x) - d_1 \frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}\delta$$

$$+\gamma f\left(x - \left(f(x) - d_1 \frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}\delta\right)$$

$$+\delta g\left(x - \left(f(x) - d_1 \frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}\delta\right)$$

Taking into account that, γ and δ respectively is a b_1 and d_1 linear translators of $f: F_{q^n} \to F_q$ and b_2 , d_2 linear translators of $g: F_{q^n} \to F_q$ we get

$$F(x)\circ G^{-1}(x)\circ H^{-1}(x)$$

$$= x - \frac{f(x)}{b_1+1}\gamma - d_1\frac{g(x)(b_1+1)-b_2f(x)}{(b_1+1)A}\cdot\gamma - \frac{g(x)(b_1+1)-b_2f(x)}{A}\delta$$

$$+\gamma f(x) - \frac{b_1 f(x)}{b_1+1}\gamma + d_1 b_1\frac{g(x)(b_1+1)-b_2f(x)}{(b_1+1)\ A}\gamma - d_1\frac{g(x)(b_1+1)-b_2f(x)}{A}\gamma$$

$$+g(x)\delta - \frac{f(x)b_2}{b_1+1}\delta + d_1 b_2\frac{g(x)(b_1+1)-b_2f(x)}{(b_1+1)A}\delta\ - d_2\frac{g(x)(b_1+1)-b_2f(x)}{A}$$

Composing similar members we have

$$= x + \gamma f(x)\left(1 - \frac{1}{b_1+1} - \frac{b_1}{b_1+1}\right) - d_1\frac{g(x)(b_1+1)-b_2f(x)}{A}\left(1 - \frac{1}{b_1+1} - \frac{b_1}{b_1+1}\right)\gamma$$

$$\frac{g(x)(b\_1+1)-b\_2\,f(x)}{A}\left(\frac{A}{b_1+1} - 1 - d_2 + \frac{d_1 b_2}{b_1+1}\right) = x$$

Theorem2

Let γ, δ, τ, ϵ F_{q^n} .Suppose γ, δ, τ, is a respectively b_1, d_1, c_1-linear translators of $f: F_{q^n} \to F_q$ and b_2, d_2, c_2 -linear translators of $g: F_{q^n} \to F_q$ and b_3, d_3, c_3 -linear translators of $l: F_{q^n} \to F_q$. Then

$$P(x) = x + \gamma f(x) + \delta\, g(x) + \tau l(x)$$

is a permutation polynomial of F_{q^n} if

1. $b_1 \neq -1$, (1)
2. $d_2 - \frac{d_1 b_2}{b_1+1} \neq -1$ (2)
3. $c_3 - \frac{b_3 c_1}{b_1+1} - \left(c_2 - \frac{b_2 c_1}{b_1+1}\right)\left(\frac{d_1 b_3 - d_3 b_1 - d_3}{(1+d_2)(b_1+1)-d_1 b_2}\right) \neq -1$ (3)

Proof

$G(x) = x + \gamma f(x)$- is a permutation polynomial in F_{q^n} by Proposition3 and condition(1).

We show that $H(x) = x + \delta\left(\frac{g(x)(b_1+1)-b_2 f(x)}{b_1+1}\right)$ is also permutation polynomial.

For convenience denote $h(x) = \frac{g(x)(b_1+1)-b_2 f(x)}{b_1+1}$.

$$h(x+\delta u) = g(x+\delta u) - \frac{b_2}{b_1+1} f(x+\delta u) = g(x) + \mathrm{d}_2 u - \frac{b_2}{b_1+1}(f(x) + \mathrm{d}_1 u) =$$

$$= h(x) + \left(\mathrm{d}_2 - \frac{d_1 b_2}{b_1+1}\right) u$$

So, δ is a $\left(\mathrm{d}_2 - \frac{d_1 b_2}{b_1+1}\right)$-linear translator of $h: F_{q^n} \to F_q$. As $\mathrm{d}_2 - \frac{\mathrm{d}_1 \mathrm{d}_2}{\mathrm{b}_1+1} \neq 0$ then according to proposition3 $H(x)$ – is also permutation polynomial in $\mathrm{F}_{\mathrm{q}^n}$.

In accordance with proposition 2

$$\mathrm{H}^{-1}(\mathrm{x}) = \mathrm{x} - \frac{\delta\ \mathrm{h}(\mathrm{x})}{1 + \mathrm{d}_2 - \frac{\mathrm{d}_1 \mathrm{b}_2}{\mathrm{b}_1+1}} = \mathrm{x} - \frac{\delta\, \mathrm{h}(\mathrm{x})}{\frac{(1+\mathrm{d}_2)(\mathrm{b}_1+1)-\mathrm{d}_1 \mathrm{b}_2}{\mathrm{b}_1+1}} = \frac{\delta\, \mathrm{h}(\mathrm{x})(\mathrm{b}_1+1)}{\mathrm{A}}$$

It is easy to see that

$$G^{-1}(x) o H^{-1}(x) = \ x - \left(f(x) - d_1 \frac{g(x)(b_1+1)-b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1)-b_2 f(x)}{A}\delta:$$

Now we consider $\mathrm{P}(\mathrm{x}) o G^{-1}(x) o H^{-1}(x)$

$$= (x + \gamma f(x) + \delta\, g(x)) o \left(x - \left(f(x) - d_1 \frac{g(x)(b_1+1)-b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1)-b_2 f(x)}{A}\delta\right)(1)$$

$$+\tau l\left(x - \left(f(x) - d_1 \frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{\gamma}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}\delta\right)$$

Since $b_1 \neq -1$ and $d_2 - \frac{d_1 b_2}{b_1+1} \neq -1$, so according to proposition 4

$$F(x) = x + \gamma f(x) + \delta\, g(x)$$

is permutation polynomial in $\mathrm{F}_{\mathrm{q}^{\mathrm{n}}}$. So by theorem1 we can imply that $(1) = x$,and we have

$$\mathrm{P(x)}oG^{-1}(x)oH^{-1}(x) =$$

$$= x + \tau\left(l(x) - \left(f(x) - d_1\frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{b_3}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}d_3\right)$$

Denote

$$l(x) - \left(f(x) - d_1\frac{g(x)(b_1+1) - b_2 f(x)}{A}\right)\frac{b_3}{b_1+1} - \frac{g(x)(b_1+1) - b_2 f(x)}{A}d_3 = k(x)\,.$$

So $\mathrm{P(x)}oG^{-1}(x)oH^{-1}(x) = x + \tau k(x)$

We show that τ is a $c_3 - \frac{b_3 c_1}{b_1+1} - \frac{1}{A}\left(c_2 - \frac{b_2 c_1}{b_1+1}\right)\left(\frac{d_1 b_3 - d_3 b_1 - d_3}{(1+\mathrm{d}_2)(\mathrm{b}_1+1) - \mathrm{d}_1\mathrm{b}_2}\right)$ linear translator of $k(x) \in F_{q^n} \to F_q$.

$$k(x+\tau u) = l(x+\tau u) - \frac{b_3}{b_1+1}\, f(x+\tau u) + \frac{d_1 b_3}{A}h(x+\tau u) - \frac{d_3(b_1+1)}{A}h(x+\tau u) =$$

$$l(x) + c_3 u - \frac{b_3}{b_1+1}(f(x) + uc_1) + \frac{d_1 b_3}{A}\left(h(x) + \left(c_2 - \frac{b_2 c_1}{b_1+1}\right)u\right)$$

$$-\frac{d_3(b_1+1)}{A}\left(h(x) + \left(c_2 - \frac{b_2 c_1}{b_1+1}\right)u\right) = l(x) + c_3 u - \frac{b_3}{b_1+1}f(x) - \frac{b_3}{b_1+1}c_1 u$$

$$+\frac{d_1 b_3}{A}h(x) + \frac{d_1 b_3}{A}\left(c_2 - \frac{b_2 c_1}{b_1+1}\right)u - \frac{d_3(b_1+1)}{A}h(x) - \frac{d_3(b_1+1)}{A}\left(c_2 - \frac{b_2 c_1}{b_1+1}\right)u$$

$$= k(x) + \left[c_3 - \frac{b_3 c_1}{b_1+1} - \frac{d_1 b_3}{(b_1+1)A}\left(c_2 - \frac{b_2 c_1}{b_1+1}\right) - \frac{d_3}{A}\left(c_2 - \frac{b_2 c_1}{b_1+1}\right)\right]u =$$

$$= k(x) + \left[c_3 - \frac{b_3 c_1}{b_1+1} - \left(c_2 - \frac{b_2 c_1}{b_1+1}\right)\left(\frac{d_1 b_3 - d_3 b_1 - d_3}{(1+\mathrm{d}_2)(\mathrm{b}_1+1) - \mathrm{d}_1\mathrm{b}_2}\right)\right]u$$

In accordance proposition3 and (3) $\mathrm{P}(\mathrm{x})oG^{-1}(x)oH^{-1}(x)$ is a permutation polynomial in $\mathrm{F}_{\mathrm{q}^{\mathrm{n}}}$. As $H(x)$ and $G(x)$ is also permutation polynomials in $\mathrm{F}_{\mathrm{q}^{\mathrm{n}}}$,then $P(x)$ also will be a permutation polynomial in $\mathrm{F}_{\mathrm{q}^{\mathrm{n}}}$.

Conclusion

In recent years in cryptography and coding theory permutations are applied very often. So it is important to propose new methods for generating permutation polynomials. Method for constructing permutation polynomials of the shape $\mathrm{P}(\mathrm{x}) = \mathrm{x} + \gamma\mathrm{f}(\mathrm{x}) + \delta\,\mathrm{g}(\mathrm{x}) + \tau\mathrm{l}(\mathrm{x})$ is given.

Bibliography

[Betti,1851] E. Betti, Sopra la risolubilit`a per radicali delle equazioni algebriche irriduttibili di grado primo, Annali di Scienze Matematiche e Fisiche 2 (1851), 5–19. (=Opere Matematiche, v.1, 17–27)

[Charpin 2009]. P.Charpin, G. Kyureghyan, When does $F(x) + \gamma Tr(H(x))$ permute F_{p^n} ?, Finite Fields .Appl15 (5) 2009 615-632

[Dickson 1896] L. E. Dickson, The analytic representation of substitutions on a power of a prime number of letters with a discussion of the linear group, Annals Math. 11 (1896-7), 65–120 and 161–183.

[Hermite 1863] Ch. Hermite, Sur les fonctions de sept lettres, C. R. Acad. Sci. Paris 57 (1863), 750–757.

[Kyureghyan G. 2011]. G. Kyureghyan, Constructing permutations of finite fields via linear translator, Journal of Combinatorial Theory 2011. P. 1052-1061

[Lidl,1987]. R. Lidl, H. Niederreiter, Finite Fields. Cambridge University Press 1987.

[Markos 2011]. J.Markos Specific permutation polynomials over finite fields Appl. 2011 V.17 p.105-112

Authors' Information

Melsik Kyureghyan *– Head of Data Coding Laboratory of Institute of Informatics and Automation problems of NAS of RA 1, P.Sevak St., Yerevan, 0014, Republic of Armenia*
e-mail: melsik@ipia.sci.am

Sergey Abrahamyan *– Researcher; Data Coding Laboratory of Institute of Informatics and Automation problems of NAS of RA 1, P.Sevak St., Yerevan, 0014, Republic of Armenia*
e-mail: serj.abrahamyan@gmail.com

SYSTEM OF INTELLIGENT SEARCH, CLASSIFICATION AND DOCUMENT SUMMARISATION FOR INTERNET PORTAL

Vyacheslav Lanin, Dmitriy Tsydvintsev

Abstract: *The article presents a description of alleged approaches to the implementation of data processing subsystem on Internet portal. Main problems are connected with exponential growth in number of documents, lack of semantic indexing and unstructured nature of information. In proposed approach, user receives an effective intelligent means of finding electronic documents on the basis of semantic indexing, automatic classification and cataloging of documents with construction of semantic links between them and automatic summarization of documents with the use of knowledge. The proposal is to increase the effectiveness of working with electronic documents with the help of intelligent analysis, for which agent-based and ontological approaches are used. In accordance with the proposed approach, ontology is used to describe data semantics of the document and its structure. Ontology is a central concept in process of document analysis. Through the use of ontologies the required data can be obtained; we know where to find information and how it can be interpreted. Ontology Repository contains three levels of ontologies. At the first level there are ontologies describing objects which are used in a particular system and which take into account system features. At the second level there are objects described in terms of the first level that are invariant to the domain. Objects of the third level describe the most general concepts and axioms, by which the lower levels objects are described. The third and second levels can be divided into two parts: a description of structures and description of the documents themselves.*

Keywords: *ontology, agent, multi-agent systems, intelligent search, semantic indexing, document analysis, adaptive information systems, CASE-technology.*

ACM Classification Keywords: *H.2. Database Management: H.2.3. Languages – Report writers; H.3.3. Information Search and Retrieval – Query formulation.*

Introduction

An exponential growth in the number of electronic documents is currently underway, and it clearly shows that traditional mechanisms for processing of electronic documents cannot cope with needs of user. This trend is evident both on Internet and in corporate networks. Currently, so-called information portals (thematic and corporate) become more and more popular, and their main objective is a consolidation of information and knowledge.

One of these solutions is a research portal – information-analytical system for the collection and analysis of data about innovation activity of regions to support effective management decisions (Research portal "Innovative development of regions"). Data for analysis is extracted from heterogeneous unstructured or semistructured data sources, in particular, Internet resources, as well as operational databases. According to the plan, the system must provide integration, coordination, aggregation and maintenance of previously disconnected data. Also it should support the various forms of data visualization and analysis, customized to the needs of users. From this it follows that the search and processing of unstructured text data from different sources in different formats, is becoming one of the main functions of the system under development.

Thus, the relevance of the problem is caused by the following reasons:

- Exponential growth in the number of documents that make it impossible to process data by traditionnal methods without loss of quality;
- Lack of semantic indexing, which does not allow for intelligent document processing in full;
- Unstructured nature of the information; the traditional mechanisms of its processing and analysis can't be used.

Consider these problems in more detail.

The exponential growth of information contained in the Internet is the reason for continuing increase in difficulty of finding relevant documents (Fig.1) and organizing them into a structured within the meaning of storage [6]. It becomes more and more difficult for user to find the necessary information; traditional search engines become ineffective.

Most technologies of working with documents focus on the organization of effective work with information for the person. But often ways to work with electronic information simply copy methods of working with paper-based information. In a text editor, there are wide range of different types of text formatting (presentation in human readable form), but little or no ability to transfer the semantic content of the text, i.e. *no semantic indexing*. To effectively address the search problem we need to expand our notion of a traditional document: *the document should be linked with knowledge to interpret and process the data stored in the document*.

Unstructured information constitutes a significant part of modern electronic documents (Fig. 2). Data Mining systems work with structured data. Unstructured content requires using Text Mining systems. In fact, they solve the same problem for different types of data, so it is assumed that these systems will converge in a "single point".

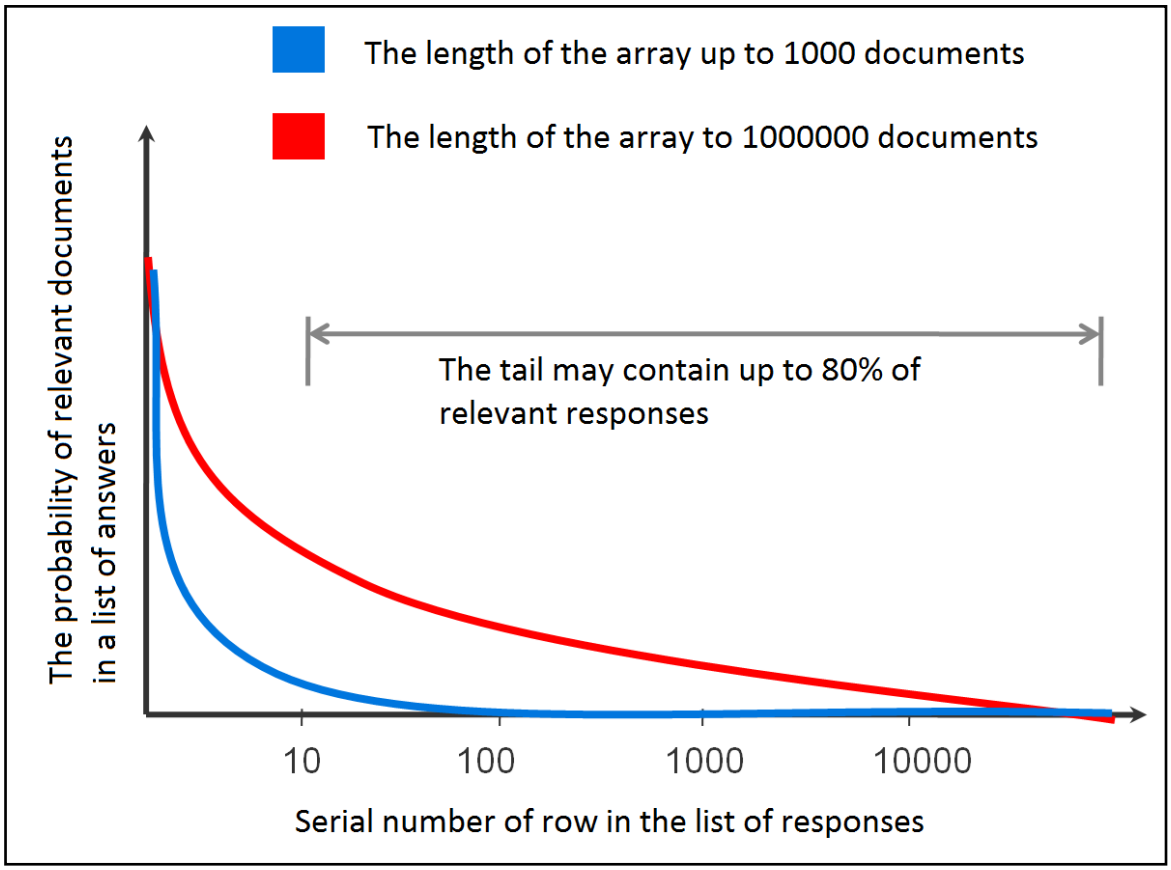


Fig.1. A problem of information retrieval with an increase in the number of documents

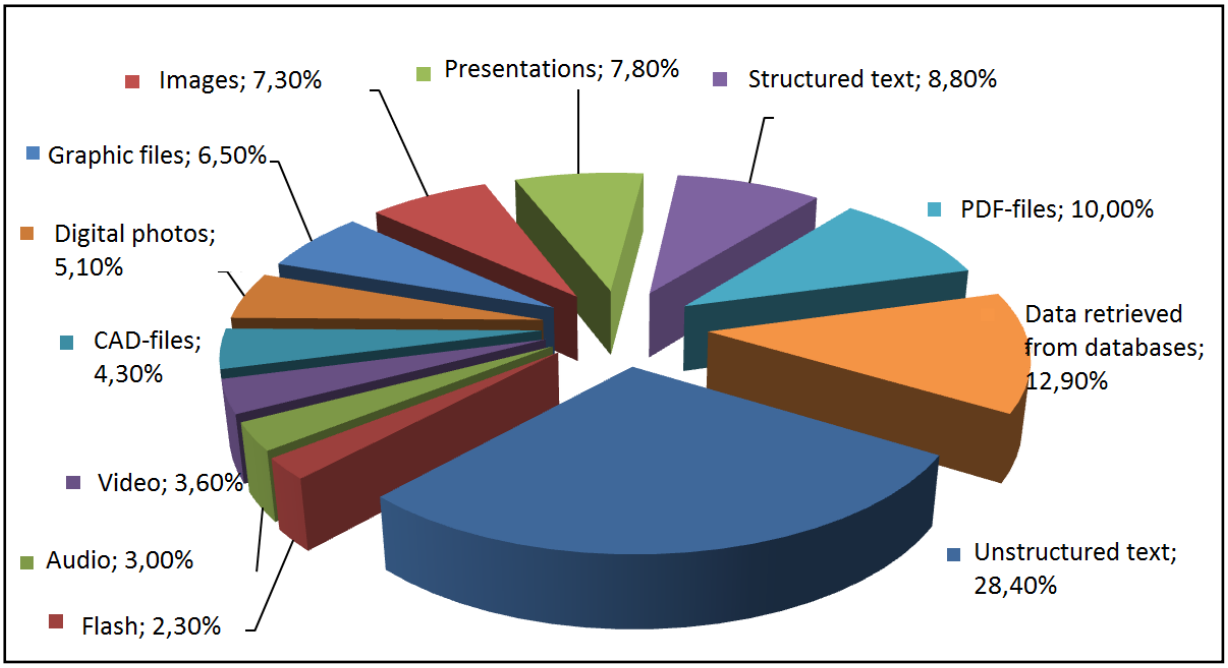


Fig. 2. Distribution of documents' categories

Text Mining allows to identify previously unknown relationships and correlations in the existing text data [5]. An important task of Text Mining technology is to derive from the text its characteristic elements or properties that can be used as document metadata, keywords or annotations. Another important task consists in assigning the document to certain categories of a given scheme of systematization. Text Mining also provides a new level of semantic search of documents. The possibilities of modern Text Mining systems can be applied in knowledge management to identify patterns in the text, to automatically "push" or post information on profiles of user interest, to create surveys of documents.

Tools and approaches of Text Mining will help to implement the intellectual capabilities of the portal when working with electronic documents.

The Approach to Semantic Indexing

Human factor has a great influence on the effectiveness of search process. User often is not ready for a long waiting of search results, for viewing and analyzing large amounts of resulting sample. In addition, most users ineffectively use search software, and usually, they ignore the advanced search capabilities, and make only short types of queries. Improvement of the electronic documents processing requires the availability of metadata describing the structure and semantics of documents. One of possible approaches to the description of information embodied in the document is an approach based on ontologies. Ontology is a knowledge base of a special type, which can be "read" and understood, alienated from the developer and / or physically separated by its users [4].

As an approach to semantic indexing has been chosen ontological approach [1], in which the ontology can describe both structure and content of the document, i.e. ontology is used to describe the semantics of the document data and its structure. Given the nature of solved targets in this paper we will concretize the notion of ontology. We assume that *ontology is a specification of certain domain*, which includes a glossary of terms (concepts) domain and a set of connections between them which describe how these terms relate to each other in a particular subject area. In fact, in this context *ontology is a hierarchical conceptual framework of the subject area*.

Ontology of document is used to analyze the document; due to it the required information can be obtained from the document: we know where to search for data and how it can be interpreted. If you represent documents using ontologies, the problem of matching ontologies and existing document is reduced to problem of search ontology terms in the document. As a result, the system needs to answer the question: Does this ontology describe the document or not. The latter question can be answered in the affirmative, if in the comparison process all the concepts included in ontology are found in document. Thus, the initial problem reduces to problem of finding general concepts in the text on the basis of formal descriptions.

Ontology repository contains *three levels of ontologies*. At the first level there are ontologies describing the objects used in a specific system and taking into account its peculiarities. The second level describes the objects that are invariant to the subject area. Objects of the second level are described in terms of objects of the first level. This is reflected in the relations of inheritance and metonymy. The objects of the third level describe the most general concepts and axioms, by which objects at the lower levels are described. The third and second levels can be divided into two parts: the description of structures and description of the documents themselves, with the documents described in terms of structures.

To address the problem of allocation of general concepts an agent-based approach is proposed on the basis of formal description [2]. This approach will satisfy the requirements of the search process, if all the advantages of multi-agent systems are realized in process of construction of the system.

When using this approach, for each node of the ontology, which contains a general concept, an agent is created which looks for this particular concept. In this approach, the agent is considered as a system aimed at achieving a particular purpose, capable of interaction with the environment and other agents. To be intelligent, the agent should have a knowledge base. Thus, to identify active agents in the system, you must choose a way to describe the knowledge base, the nature of interaction with the environment and cooperation.

Knowledge base of agent for finding find common concepts of the ontology can be also conveniently presented in the form of ontology. To enable the user to add new templates it is necessary to select the basic concepts for the formation of general ones.

One of the most important properties of agents is *sociality*, or the ability to interact [2]. As mentioned above, the agent is created for each node of the ontology, which contains a general concept. According to the accepted classification of agents it is *intentional agent*.

This agent is designed to address two problems:

1. It breaks the entire list of available templates concepts into separate components and runs simple search agents for searching of derived components.
2. Assembles results from all the lists submitted by agents of the lower level.

The agents at a lower level mentioned above are called *reflex agents*. They get a template, and their goal becomes finding phrases in the text covered by this template. Search results for agents of all levels shall be recorded on the "bulletin board".

At this point in other systems the instruments of ontological nature are used in the following areas:

- WordNet in conjunction with the vector and Boolean models of information retrieval;
- Traditional information retrieval thesauri in combination with various statistical models;
- Thesaurus for automatic indexing in Boolean models of documents searching, in problem of automatic headings and automatic annotation.

Ontologies will form a core of portal metadata when working with electronic documents. Clearly defined subject area allows creating sufficiently detailed ontologies, which can be used by all its subsystems.

Automatic Abstracting

Currently, there are two approaches used for automatic summarization. A traditional approach (quasi abstracting), which is used by such systems as Microsoft Office, IBM Intelligent Text Miner, Oracle Context, is based on allocation and selection of text fragments from the source document and connection them in a short text. On the other side, there is approach based on knowledge, which involves preparation of summaries and transfers the basic idea of the text, perhaps even in other words.

Quasi abstracting is based on the allocation of specific fragments (usually sentences). For this purpose, a method of comparison of phrasal templates chooses blocks with the greatest lexical and statistical relevance. There are a model of linear weights is used in most implementations of the method. The analytical phase of this model is a procedure of appointing the weighting coefficients for each block of text in accordance with such characteristics as location of this block in the original, frequency of appearance in the text, frequency of use in key proposals, as well as indicators of statistical significance. So, there are three main directions, often used in combination: statistical methods, positional methods and indicator methods.

The main advantage of this model lies in the simplicity of its implementation. However, the selection of sentences or paragraphs, not taking into account the relationship between them, leads to the formation of disconnected essays. Some proposals may be omitted, or there can be "hanging" words or phrases in them.

To implement the second method, some ontological reference is needed, reflecting the views of common sense and the concepts of targeted subject area, to make decisions during the analysis and to determine the most important information.

Method of forming a summary suggests two basic approaches.

The first approach relies on the traditional linguistic method of parsing sentences. This method is also uses semantic information to annotate parse trees. Comparing procedures directly manipulate the trees to remove and rearrange the parts, for example, by reducing the branches on the basis of certain structural criteria, such as brackets or embedded conditional or subordinate sentences. After this procedure, the parse tree is greatly simplified, becoming, in essence, a structural "squeeze" of the original text.

The second approach for compiling a summary roots in artificial intelligence systems and relies on natural language understanding. Parsing is also part of such a method of analysis, but the parse tree in this case is not generated. On the contrary, there are conceptual representative structures of all the initial information are formed, which accumulate in the text knowledge base. As the structures, formulas of predicate logic or such representations as a semantic network or frame set can be used.

Automatic summarization is necessary for the developed portal. When user is searching, it is necessary to present him a document annotation, by which he can decide on the usefulness of this document.

Classification and Cataloging of Documents

The task of automatic classification and cataloging of documents is the task of partitioning the incoming stream of text into thematic substreams according to predetermined headings. Automatic cataloging of electronic documents, and documents posted on Internet in particular, is complicated because of the following reasons [8]:

- A large array of documents;
- An absence of special structures for tracking the emergence of new documents;
- Optionality of the author's classification of electronic documents (as opposed to print publications) through annotation, attribution of the qualifier codes, etc.;
- A problem of tracking changes in documents.

As for automatic abstracting, there are two opposite approaches to cataloging. *The methods based on knowledge* are the most effective, but it's difficult to implement them. When cataloging the texts on the basis of knowledge preformed knowledge bases are used. They describe language expressions, corresponding to a particular category, and rules for the selection of headings [5]. Another class of methods for automatic categorization of texts is *the methods of machine learning*, which can use manually pre-cataloged texts as training examples.

When implementing a system of automatic cataloging of the portal, it is necessary to solve two problems:

- *Establishment of a mechanism for introduction and description of categories*, as some expression on the basis of words and terms in documents. The problem can be solved on the basis of expert descriptions of categories or on the basis of machine learning methods with the help of pre-cataloged collections of documents.
- *Analysis of linguistic material and context of words' using*. It requires an extensive knowledge of the language and subject area.

Conclusion

The above approaches are used in the development of an electronic document management subsystem of research portal. Its distinctive feature is focus on the explicit knowledge representation by using ontologies. This approach will allow us to realize intelligent services for searching and processing of electronic documents related to the portal and gathered from different sources.

As a result, the following tasks will be solved by creating the research portal:

- Semantic indexing of documents and intelligent retrieval of data corresponding to users' queries and the specific subject area;
- Extracting information from unstructured documents;

– Intellectual classification and cataloging and automatic summarization of retrieved documents;
– Maintaining a history of electronic documents.

The implementation of the subsystem will significantly reduce complexity to find useful information, its analysis and possible use in research.

References

[1] Lanin V. Intelligent management of documents as the basis for the technology of adaptive information systems // Proceedings of the International Scientific-Technical Conference «Intelligent systems» (AIS'07). V. 2 / M.: Fizmatlit, 2007. P. 334-339.

[2] Tarasov V. From multi-agent systems to intelligent organizations: philosophy, psychology, computer science. M.: Editorial, URSS, 2002.

[4] Khoroshevskii V., Gavrilova T. Knowledge Base intelligent systems. Petersburg.: Peter, 2001.

[5] Lande D. Knowledge Search on the Internet. Professional work. M.: Publishing house "Williams", 2005.

[6] Efremov V. Search 2.0: fire on the "tail" // Open systems. DBMS № 08 (134), 2007.

[7] Chernyak L. Enterprise Search 2.0 // Database. - 2007. - № 07 (133).

[8] Fedotov A, Barakhnin V. Internet Resources as an object of scientific research [electronic resource. - 2007. - Mode of access: http://www.rfbr.ru/pics/28320ref/file.pdf.

[9] Weal M.J., Kim S., Lewis P.H., Millard D.E., Sinclair P.A.S., De Roure D.C., Nigel R. Ontologies as facilitators for repurposing web documents / Shadbolt. Southampton, 2007.

Authors Information

Vyacheslav Lanin *– Perm State University, Department of Computer Science; Russia, Perm, 614990, Bukirev St.,15; e-mail: lanin@psu.ru.*

Dmitriy Tsydvintsev *– Perm State University, Department of Computer Science; Russia, Perm, 614990, Bukirev St.,15; e-mail: akinokinos@yandex.ru.*

WAVELET TRANSFORM IN INVESTIGATIONS OF STUDENTS EDUCABILITY DEPENDENTLY ON DEGREE OF GRAPHICAL SKILLS AUTOMATION OF WRITING PROCESS

Olga Kozina, Nataliya Bilous, Mykola Zapolovskij, Vladimir Panchenko

Abstract: Application of 2-level 2-D Haar Discrete Wavelet Transform of images with special words repeated by hand is offered. Metric of similarity for images with pattern words and images with students' written words like analog of proposed patterns is considered. A new approach to objective assessment of educability like dependence on level of graphic skills of process of writing is offered. Experiment is revealed the presence of statistic significant correlation between average score in special academic subjects group or in other words between special educability and level of graphic writing skills automation for high school students.

Keywords: educability, process of writing, 2-D discrete wavelet transform, image similarity

***ACM Classification Keywords**: I.5 PATTERN RECOGNITION and K.3.2 COMPUTER AND INFORMATION SCIENCE EDUCATION*

1 The degree of automation of movements in writing process like a feature of the higher nervous activity

Process of writing is a complex activity and all divisions of the cerebral cortex participate in its formation. Psychophysical basis of process of writing is the interaction of different analyzers – speech-motor, auditory, visual, hand-motor. The interaction of such mental activities as thinking, memory, attention, imagination, external and internal speech occurs during writing. Process of writing is a complex motor skill, which includes technical, graphical and spelling skills. Graphic skills reflect the ability to write symbols by hand on paper quickly and clearly and spelling skills reflect knowledge how to apply hand-written characters and the abilities to use these rules during process of writing for the reflecting of its content. Technical skill is manifested in the ability to hold the right techniques and methods of writing, such as proper position of the body during process of writing or the location of paper, etc. The technical and graphic skills define a character and features of the handwriting of a person.

The process of writing according to its psychological content from the outset is a conscious act to be formed in the special education arbitrarily and requires a number of special operations for its implementation in contrary to spoken language which assimilated across the imitating of another person's speech.

The type of coordination of movements during process of writing that acquired in learning and the degree of automation of these movements greatly influence on the number and nature of deviations appeared during

writing, from the typical patterns of writing of characters and symbols. Insufficient degree of automation of movements during process of writing will be manifested in significant changes in the implementation of characters and their compounds than it is happening in the cases with a sufficient degree of automation of these movements [Kornev 1999, Bernstein 1990].

Processes of excitations and inhibitions are main nervous processes taking place in development of writing. Ratio of the basic properties of these processes (power, balance, mobility) affects the formation of writing. Type of higher nervous activity, depending on the degree of dominance first or second signaling system in it and the degree of precision transfer of temporal links the first signaling system in the second one, which may be different for individuals, may affect the ability of that person more or less successfully imitate of handwriting features of another person. I. Pavlov identifies three main types of people depending on the predominance of the first or second signal system in the brain [Bernstein 1990]:

a) the artistic style, characterized by a predominance of the first signaling system;

b) the thinking style, characterized by a predominance of the second signaling system;

c) the average type, characterized by a relative balance of both signaling systems.

Other things being equal, person with a predominance of the first signaling system in the state more accurately imitate the handwriting of another person and demonstrate in his process of writing to the significant set of handwriting features of another person than those with a predominance of the second signal system. It is known that persons with a more precision transfer of temporal links the first signaling system in the second signaling system able to fully and comprehensively recognize the signs of his and others' handwriting and successfully imitate the handwriting of another person [Kornev 1999].

Thus, individual differences in process of writing to different persons are a reflection not only of external learning environments, but also characteristics of their higher nervous activity.

2 organization of the experiment to identify the statistical relationship between educability and the level of graphical skills automation of writing process

Pavlov's research showed that education is the formation of temporary connections in the cerebral cortex of the brain, and skill – the system of these bonds [Podlasiy 2003].

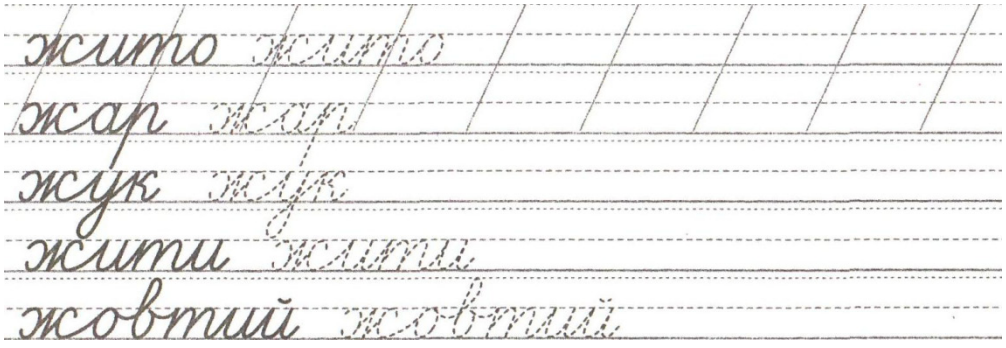
On the one hand, the motor component of process of writing requires considerable conscious effort and attention when writing, distracting him from more important tasks associated with given semantic task. On the other hand, it's known that productivity of education is directly proportional to level of educability and stability of attention of students, and it also depends on the level of their memory. Addition, the student achievements are directly proportional to educability depending on the strength of attention. Therefore, if we assume that student estimations reflect student's achievements and student with higher scores has greater educability, then the level

of automation of the motor components of process of writing is a reflection of the level of attention, therefore, the level of educability that is the main hypothesis will be tested experimental.

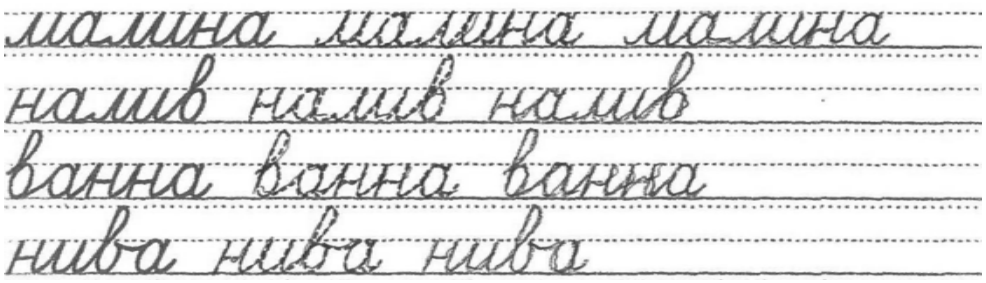
Distinguish between general educability – the ability of mastering any material and special educability – learning ability of certain types of material (of various sciences, academic subjects, practices). The first type of educability is an indicator of the total talents of the individual while the second one is an indicator of only special.

To identify the statistical relationship between each type of educability and the level of graphical skills of writing process works from 18 third-year students of Computer and Information Technologies Faculty in the National technical university «Kharkiv Politechnical Institute» (Ukraine) were analyzed. Forms with patterns of words with different lengths in native language were given for students in several weeks. One part of such sheets is shown in Figure1a.

Spelling and fonts of handwriting patterns are conform to the norms of the Ministry of Education and Science of Ukraine and were taken from handwriting patterns for primary classes. Students were not aware of the purpose of the experiment, in other words they were not aware that the quality of this works will be compared with their educability. For students the problem had formulated such: as precisely as possible to repeat the handwriting patterns of words and to continue write to string the same words in the shortest time. Then an average score was calculated from previous 4 sessions for each student. Moreover, all subjects were divided into two groups – the general and special subjects. Each group of academic subjects is corresponds to general or special educability.



a



b

Fig.1. Examples of experimental materials: a – part of a sheet-pattern, b – part of a sheet-result

Table 1. Groups of academic subjects

General subjects	Special subjects
High Mathematics	Programming
Ukraine History	Basics of Computer Technology
Physics	Computer-aided design
Engineering Graphics	Discrete Mathematics
Theory of electro-magnetic circuit	System Programming
Philosophy. Logic	Applied theory of digital automata
Foreign Language	Computer electronics and circuitry
History of Science and Technology	Computing Algorithms and Data Structures
Cultural Studies, Ethics , Religious studies	Databases
Sport	Computers Architecture

Thus, the level of automation of graphical and motor skills of process of writing reinforced by the ability to properly carry out the assigned trivial tasks are compared with two values of average scores on general and special academic subjects.

3 Processing of experimental data

20 repeated and 20 own written three-, four- and five-symbols words for each student were selected from all obtained materials. To assess the level of graphical skills automation of writing process measure of similarity between the proposed handwriting patterns of word and both repeated word and own written same word are analyzed. Digital variants of experiment materials was received by scanning of the completed student sheets with experimental results and clean patterns sheets with same resolution. Algorithm for primary processing of digital images has been developed and implemented in Matlab. The sense of the algorithm is as follows:

- alignment of contours on patterns sheets and results sheets;

- subdividing of whole sheets on disjoint sub-images with patterns of every words;

- select a base point on the processed results sheet which will be left down vertex of the allocated area of interest with equal size dimensions to pattern sub-image of corresponding word (Figure 2);

- convert the result sub-image with analyzed word into grayscale.

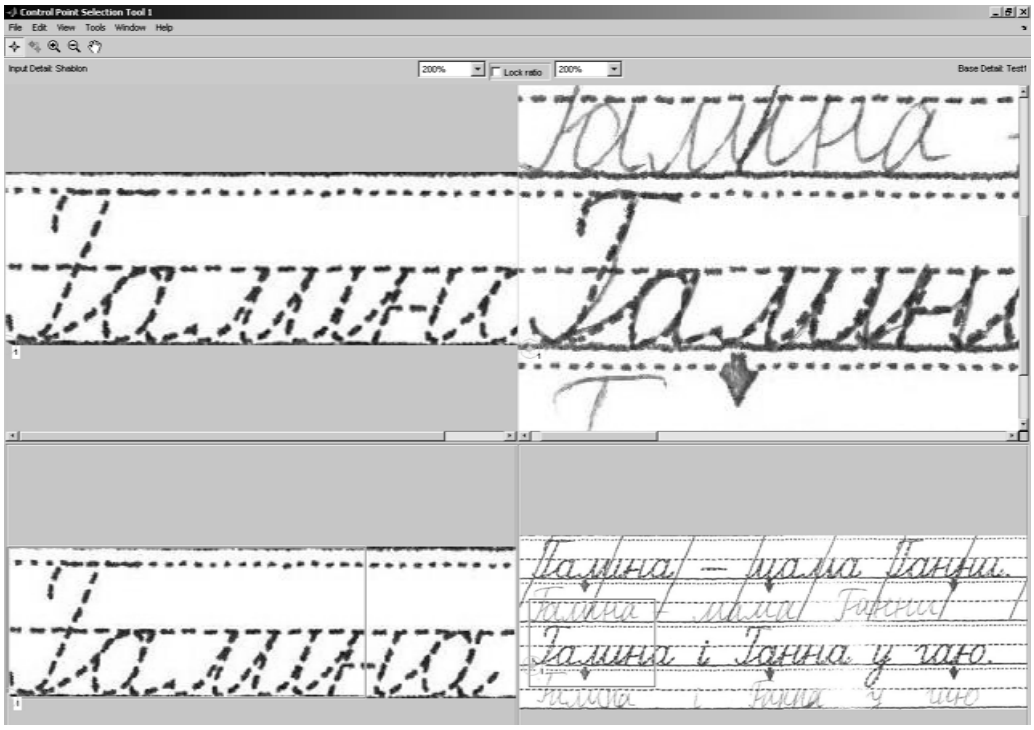


Fig. 2. Selection of interest area into results sheet

After primary processing the task of experiment results processing are considered as a problem of quantitative comparison of pattern and result sub-images with corresponding word through a specific metric called similarity measurement. A high similarity value between two images means a high level of graphical skills automation of writing process.

To images recognizing and to their similarity estimation Fourier Transform and discrete wavelet transform are successfully used [Bebis 2006, Pretto 2010, Rakhmankulov 2008].

As presented in [Burrus 1998] the Fourier Transform gives the spectral content of the whole signal, but it gives no explicit information regarding where in space those spectral components appear. A better tool for non-stationary signal analysis (whose frequency response varies in time, like in the images) is the Wavelet Transform [Iyengar 1997, Chui 1992, Mallat 1998]: it gives information about which frequency components exist and where these components appear. Properties of Discrete Wavelet Transform (DWT) coefficients were exploited in order to

calculate the images similarity for our case. Every image can be represented as a linear combination of the images in the wavelet basis:

$$P(x) = \sum_k c_k \varphi_k(x) + \sum_k \sum_j d_{j,k} \psi_{j,k}(x) \quad (1)$$

where c_k – approximation coefficients, $d_{j,k}$ – detailed coefficients, k – location index, j – scale parameter.

A wavelet representation of a function consists of a coarse overall approximation together with detail coefficients that influence the function at various scales. For our experiment 2-D Haar DWT of the grayscale images was selected. Given the resolution of images, we decide to stop at second level of decomposition.

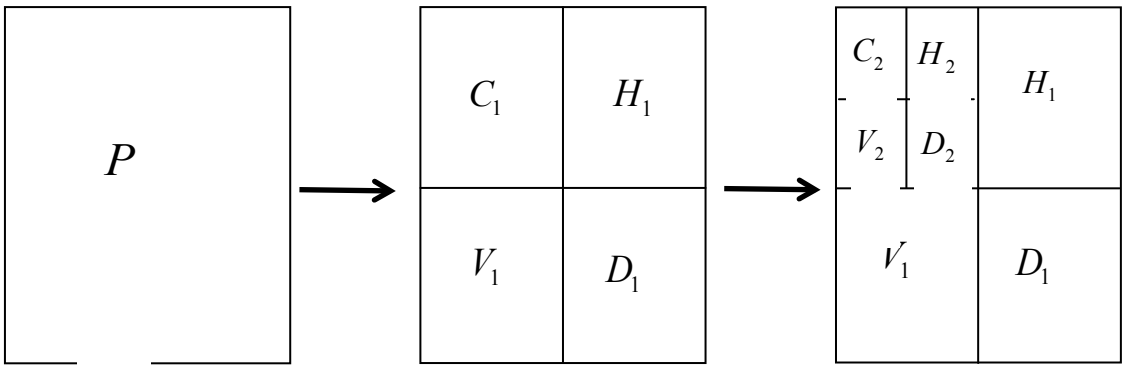


Fig.3. Two-level 2-D Wavelet decomposition

2-level 2-D Wavelet decomposition of input image P_i, and C_j – the approximation coefficients, H_j, V_j, D_j – respectively the horizontal, vertical and diagonal detailed coefficients in *j*-th level of the wavelet decomposition are shown in Figure 3. Transform going to higher level decompositions can significantly reduce the decomposed image's size, but higher level of decompositions discard important features in images, like edges and high frequency patterns useful for comparing of images in our experiment.

The Haar Wavelet is chosen as a wavelet type because of it is very effective in detecting the exact locations when a signal changes: image discontinuity is one of most important features chosen in image-based localization. The detail coefficients are the intensity variations along rows, columns and diagonals. Thus detailed coefficients can be used to detect and highlight image shapes. The higher the absolute values of coefficients are the higher probabilities that it encodes salient image features.

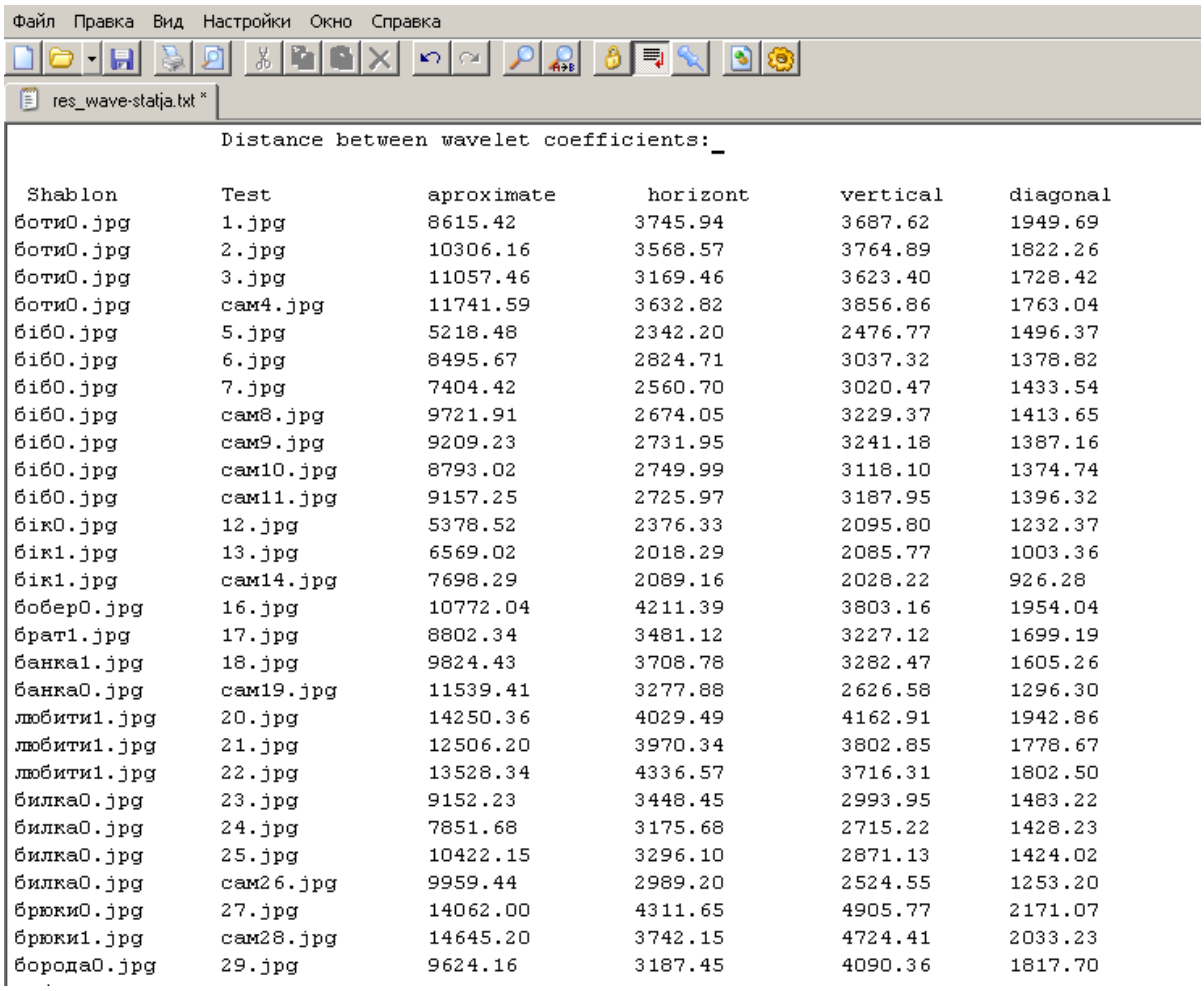
2 sets of 4 matrices of the horizontal, vertical and diagonal detailed coefficients with dimensions $a \times b = w/2^{wavelet....level} \times h/2^{wavelet....level}$ where w is the image's width and h is the image's height were obtained by using the functions from Matlab to calculate the coefficients of 2-level 2-D Haar DWT for each pair of pattern-result sub-images.

A metric that compares how many significant wavelet coefficients the query has in common with potential targets is used to compute image similarity usually [Pretto 2010]. But in our case, such a metric is not suitable, especially for the analysis of the repeated handwriting words. Therefore, considering the obtained matrix with coefficients as coordinates of multidimensional space, calculate the length of each vector $\rho(C_2), \rho(H_2), \rho(V_2), \rho(D_2)$ by the formula calculating the Euclidean distance between the coordinates of wavelet coefficients space for pattern and result sub-images with corresponding word:

$$\begin{aligned}
\rho(C_2) &= \sqrt{\sum_{l=1}^{a}\left(C_2^{sh}(l) - C_2^{t}(l)\right)^2 + \sum_{c=1}^{b}\left(C_2^{sh}(c) - C_2^{t}(c)\right)^2}, \\
\rho(H_2) &= \sqrt{\sum_{l=1}^{a}\left(H_2^{sh}(l) - H_2^{t}(l)\right)^2 + \sum_{c=1}^{b}\left(H_2^{sh}(c) - H_2^{t}(c)\right)^2}, \\
\rho(V_2) &= \sqrt{\sum_{l=1}^{a}\left(V_2^{sh}(l) - V_2^{t}(l)\right)^2 + \sum_{c=1}^{b}\left(V_2^{sh}(c) - V_2^{t}(c)\right)^2}, \\
\rho(D_2) &= \sqrt{\sum_{l=1}^{a}\left(D_2^{sh}(l) - D_2^{t}(l)\right)^2 + \sum_{c=1}^{b}\left(D_2^{sh}(c) - D_2^{t}(c)\right)^2}
\end{aligned} \tag{2}$$

where *sh* – index of matrixes' elements for approximation coefficients, for horizontal, diagonal and vertical detailed coefficients of second level wavelet decomposition of pattern sub-image, respectively,

t – index of matrixes elements for approximation coefficients, for horizontal, diagonal and vertical detailed coefficients of second level wavelet decomposition of result sub-image, respectively.



Distance between wavelet coefficients:

Shablon	Test	aproximate	horizont	vertical	diagonal
ботиО.jpg	1.jpg	8615.42	3745.94	3687.62	1949.69
ботиО.jpg	2.jpg	10306.16	3568.57	3764.89	1822.26
ботиО.jpg	3.jpg	11057.46	3169.46	3623.40	1728.42
ботиО.jpg	сам4.jpg	11741.59	3632.82	3856.86	1763.04
бібО.jpg	5.jpg	5218.48	2342.20	2476.77	1496.37
бібО.jpg	6.jpg	8495.67	2824.71	3037.32	1378.82
бібО.jpg	7.jpg	7404.42	2560.70	3020.47	1433.54
бібО.jpg	сам8.jpg	9721.91	2674.05	3229.37	1413.65
бібО.jpg	сам9.jpg	9209.23	2731.95	3241.18	1387.16
бібО.jpg	сам10.jpg	8793.02	2749.99	3118.10	1374.74
бібО.jpg	сам11.jpg	9157.25	2725.97	3187.95	1396.32
бікО.jpg	12.jpg	5378.52	2376.33	2095.80	1232.37
бік1.jpg	13.jpg	6569.02	2018.29	2085.77	1003.36
бік1.jpg	сам14.jpg	7698.29	2089.16	2028.22	926.28
боберО.jpg	16.jpg	10772.04	4211.39	3803.16	1954.04
брат1.jpg	17.jpg	8802.34	3481.12	3227.12	1699.19
банка1.jpg	18.jpg	9824.43	3708.78	3282.47	1605.26
банкаО.jpg	сам19.jpg	11539.41	3277.88	2626.58	1296.30
любити1.jpg	20.jpg	14250.36	4029.49	4162.91	1942.86
любити1.jpg	21.jpg	12506.20	3970.34	3802.85	1778.67
любити1.jpg	22.jpg	13528.34	4336.57	3716.31	1802.50
билкаО.jpg	23.jpg	9152.23	3448.45	2993.95	1483.22
билкаО.jpg	24.jpg	7851.68	3175.68	2715.22	1428.23
билкаО.jpg	25.jpg	10422.15	3296.10	2871.13	1424.02
билкаО.jpg	сам26.jpg	9959.44	2989.20	2524.55	1253.20
брюкиО.jpg	27.jpg	14062.00	4311.65	4905.77	2171.07
брюки1.jpg	сам28.jpg	14645.20	3742.15	4724.41	2033.23
бородаО.jpg	29.jpg	9624.16	3187.45	4090.36	1817.70

Fig.4. Example of part of results sub-images processing for one student

An example of the file with obtained according to (2) distances $\rho(C_2), \rho(H_2), \rho(V_2), \rho(D_2)$ between all pattern and result sub-images accordingly for one student are shown on Figure 4. Left two columns contain the names of files with compared sub-images and the next 4 columns contain $\rho(C_2), \rho(H_2), \rho(V_2), \rho(D_2)$ respectively. Obtained distances were grouped in 3 samples, depending on the length of the word-pattern. For each sample were calculated standard deviations.

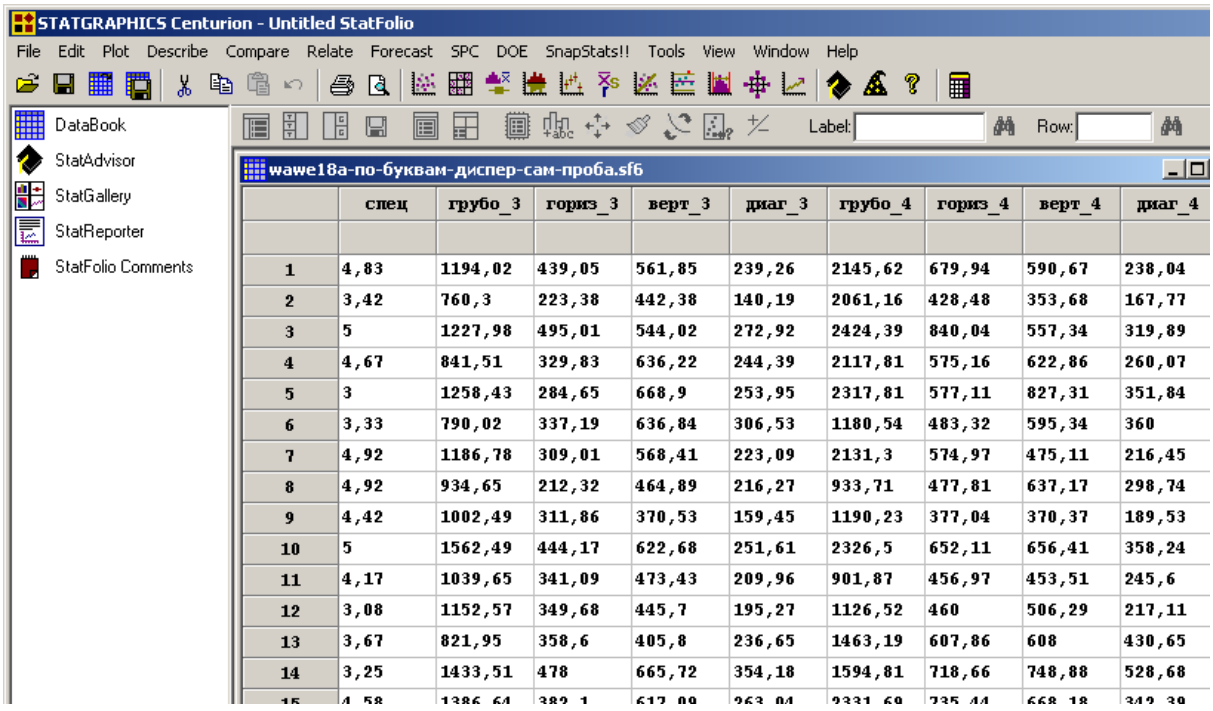
Standard deviation of the distance between the corresponding wavelet coefficients of pattern and result sub-images was calculated for each length of word. Correlation coefficients between the average score of general and special subjects groups disjoint and standard deviations of distances between 2D Haar DWT coefficients for every lengths of word were calculated in Statgraphics Centurion package.

Each line in Figure 5 contains data of one student in the following sequence: average scores on special academic subjects, standard deviations of $\rho(C_2), \rho(H_2), \rho(V_2), \rho(D_2)$ for three-symbol words and standard deviations of $\rho(C_2), \rho(H_2), \rho(V_2), \rho(D_2)$ for four-symbol words.

Pearson product moment correlations between each pair of variables x and y calculated by:

$$r(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (3)$$

are represented in Table 2. P-values below 0,05 indicate statistically significant non-zero correlations at the 95,0% confidence level.



	спец	грубо_3	гориз_3	верт_3	диаг_3	грубо_4	гориз_4	верт_4	диаг_4
1	4,83	1194,02	439,05	561,85	239,26	2145,62	679,94	590,67	238,04
2	3,42	760,3	223,38	442,38	140,19	2061,16	428,48	353,68	167,77
3	5	1227,98	495,01	544,02	272,92	2424,39	840,04	557,34	319,89
4	4,67	841,51	329,83	636,22	244,39	2117,81	575,16	622,86	260,07
5	3	1258,43	284,65	668,9	253,95	2317,81	577,11	827,31	351,84
6	3,33	790,02	337,19	636,84	306,53	1180,54	483,32	595,34	360
7	4,92	1186,78	309,01	568,41	223,09	2131,3	574,97	475,11	216,45
8	4,92	934,65	212,32	464,89	216,27	933,71	477,81	637,17	298,74
9	4,42	1002,49	311,86	370,53	159,45	1190,23	377,04	370,37	189,53
10	5	1562,49	444,17	622,68	251,61	2326,5	652,11	656,41	358,24
11	4,17	1039,65	341,09	473,43	209,96	901,87	456,97	453,51	245,6
12	3,08	1152,57	349,68	445,7	195,27	1126,52	460	506,29	217,11
13	3,67	821,95	358,6	405,8	236,65	1463,19	607,86	608	430,65
14	3,25	1433,51	478	665,72	354,18	1594,81	718,66	748,88	528,68
15	4,58	1386,64	382,1	617,09	263,04	2331,69	735,44	668,18	342,39

Fig.5. Initial data for calculations of correlation coefficients

Table 2. Correlation coefficients between variables x and y

Number of symbols into word	Correlation coefficients $r(x,y)$				
	x	Group of academic subjects or educability type			
	Type of vector	y special	P-value	y general	P-value
3	$\rho(C_2)$	0,4752	0,0539	0,3952	0,1164
	$\rho(H_2)+\rho(V_2)+\rho(D_2)$	0,1459	0,5764	0,0382	0,8842
4	$\rho(C_2)$	**0,5552**	0,0207	**0,6807**	0,0026
	$\rho(H_2)$	**0,5318**	0,0280	0,4174	0,0955
5	$\rho(C_2)$	**0,6905**	0,0022	**0,6228**	0,0076
	$\rho(H_2)+\rho(V_2)+\rho(D_2)$	0,3418	0,1793	**0,5378**	0,0260

4 Interpretation of results and future works

Considering the Haar Wavelet, the approximation coefficient C_2 represent only the mean of the intensity of the pixels composing the macro-squares (4x4 pixels in our case) or general characteristic of a word. On the other hand, the difference between values of this coefficient for pattern and result sub-images is proportional to the number of pixels which involved to writing of analog of the pattern word but have distinct intensities of pixels at the pattern word or to general recognition, the readability of the word.

The horizontal, vertical and diagonal detailed coefficients H_2, V_2, D_2 are shown preferential direction of inaccuracies in the writing of words. Since the maximum value correlation coefficient between $\rho(H_2)$ and the average score in special academic subjects group for 4-synbols words (r = 0.5318) with no statistically significant correlation with $\rho(V_2)$ (r = 0.2389 and P-value = 0.3558) demonstrates that the loss of attention during process of writing leads to more extreme variations in the width of written word than variations of its height .

Absence statistically significant correlation coefficients between the mean scores for both groups of subjects and three-symbols words requires further study and may be explained by the fact that when writing of short words "truncated" motor skills are predominated in nervous activities of students, similar to the graphical skills of alone symbol writing. Movements of hand in such cases are minimum also and do not require a long and precise control of motor skills. Besides shorter words easier to remember, that is why smaller loss of attention occurs during process of writing short words. This gives rise to the formation of hypotheses about the higher values of correlation coefficients between special educability and measures of patterns-results sub-images similarity for seven-, eight- or nine-symbols words. Verifying this assumption, as well as the formulation of regression models of students' educability relating to the level of graphical skills automation of students will be the subject of future studies.

Conclusion

Thus, described experiments revealed the presence of statistic significant correlation between average score in special academic subjects group or in other words between special educability and level of graphic writing skills automation for high school students. Further study about abilities of using 2-D Discrete Wavelet Transform to quantify the level of graphic skills automation of writing process should give base for models of an objective assessment of educability.

Bibliography

[Kornev 1999] A. Kornev, Methodology to evaluate of automation of writing skills, Problems of pathology of development and disintegration of speech function. St. Petersburg: Pub. of St.-Petersburg University (1999), pp.54–59 (in Russian).

[Bernstein 1990] N. Bernstein, Physiology of movements and activity, Moscow: Nauka, (1990), 494p (in Russian).

[Podlasiy 2003] I. Podlasiy, Effective Pedagogy: the book for teachers, Moscow: Public education, (2003), 496p (in Russian).

[Bebis 2006] G. Bebis, A. Gyaourova, S. Singh, I. Pavlidis, Face Recognition by Fusing Thermal Infrared and Visible Imagery, Image and Vision Computing, vol. 24, no. 7 (2006), pp. 727-742.

[Pretto 2010] A. Pretto, E. Menegatti, Y. Jitsukawa, R. Ueda, T. Arai, Image similarity based on Discrete Wavelet Transform for robots with low-computational resources, Robotics and Autonomous Systems, 58 (2010) 879–888pp.

[Rakhmankulov 2008] V. Rakhmankulov, A. Akhrem, V. Gerasimova, A. Barsegov, Multidimensional wavelet analysis of images, Proceedings of ISA RAS T., 38 (2008), pp.278-288 (in Russian).

[Burrus 1998] S. Burrus, R. Gopinath, H. Guo, Introduction to Wavelets and Wavelet Transforms: A Primer, New Jersey: Prentice Hall, (1998), 282 pp.

[Iyengar 1997] S. Iyengar, L. Prasad, Wavelet Analysis with an Application to Image Processing, Chapman and Hall/CRC Press, (1997), 279p.

[Chui 1992] C. Chui, An Introduction to Wavelets, London: Academic Press, (1992), 266p.

[Mallat 1998] S. Mallat, A Wavelet Tour of Signal Processing, New York: Academic Press, (1998), 637p.

Authors' Information

Olga Kozina *– Ph.D., lecturer of Computers and Programming Department of National Technical University 'KPI', Kharkiv, Ukraine; e-mail: hpi-kozina@rambler.ru*

Major Fields of Scientific Research: medical information systems, segmentation and recognition ,system modeling

Nataliya Bilous *– Ph.D., head of ITLCVS laboratory, professor of Kharkov National University of Radio Electronics, Kharkov, Ukraine; e-mail: belous@kture.kharkov.ua*

Major Fields of Scientific Research: image processing, segmentation and recognition, building of computer vision system, computer testing and learning systems

Nikolaj Zapolovskij *– Ph.D., Cheaf of Computer and Information Technologies Faculty of National Technical University 'KPI', professor of National Technical University 'KPI', Kharkiv, Ukraine*

Major Fields of Scientific Research:system modeling, computer testing and learning systems

Vladimir Panchenko *– teacher of Computers and Programming Department of National Technical University 'KPI', Kharkiv, Ukraine*

Major Fields of Scientific Research: synthesis of tests for learning systems, artificial intelligence systems

A COMPUTER METHOD TO STUDY THE ENTIRETY OF STUDENTS' KNOWLEDGE ACQUIRED DURING AN EDUCATIONAL COURSE

Evgeny A. Eremin

Abstract*: This paper considers the experimental research of measuring interconnections between the basic concepts, acquired during completion of a course in Computer Architecture. A special computer technique for estimating of students' knowledge entirety was developed and successfully tested. An original mode of experimental data visualization is proposed. Several pedagogical regularities were revealed by assessing how students digested the main concepts of the course.*

Keywords*: entirety, concept, knowledge structure, education, course.*

ACM Classification Keywords*: K.3.1 Computer Uses in Education; I.2.6 Learning – Knowledge acquisition.*

Introduction

Methodology and consistency are always noted as the chief mainframe principles of education in all pedagogical books from classics to modern manuals. «A man has veritable and real knowledge only when clear-cut picture of the outer world, which is represented as a system of interrelated concepts, is reflected in his brain; … if you do not keep up methodology and consistency in education, the process of pupil's cultivation slows down.» [Podlasiy, 2002] It is emphasized, that «formed system of knowledge is the most important recipe to prevent its loss. Forgotten knowledge is quickly restored in a system, and hardly without it. … Do not forget J.A. Komensky's advice: all must be pursued in an indissoluble sequence, so that all today's knowledge confirms yesterday's and paves the way for tomorrow.»

The well known Russian lecturer A. Kushnirenko, who is the author of several informatics textbooks, generalized the idea about necessity of learning process's unity, using the following capacious and exact thesis: pupils must «digest such minimum of knowledge, accumulated by the mankind, **that permits shaping the entire representation of outward things, nature and society** » [Kushnirenko , 1998].

Unity of interrelated concepts in any concrete educational course in many respects determinates the success of learning, and the entirety of educational material is always identified as a mandatory pedagogical requirement. In a publication [Zorina, 1992], which analyses the didactic problems of selecting knowledge to include in a

textbook, the entirety of course's contents is marked out as the second principle by its significance (straight after the necessity of science matter reflected in a learning subject). More detailed citation is the following: «Constructing a scientific content of the educational discipline we need to orient oneself to its entire reflection in a textbook. … The entirety of the whole course realizes itself in a scientific picture of the world. The entirety of reflecting different elements of knowledge realizes through their constitution and structure.»

To avoid ambiguity in the discussion, let us arrange that the term «**entirety of the concepts' system**, built for the learning course», means the presence of the essential relations between all basic categories of this system. As follows from the proposed statement, those students, who see more associations between studied terms, have a higher rate of knowledge entirety, and, hence, digested this learning course better.

This paper is an attempt of experimental research, aimed at studying the structure of concepts interrelations, which was formed as a result of the course «Computer Architecture». The special computer method was developed and tested in order to estimate the rate of entirety of the system from basic concepts, which students acquired by studying the material. Several gauges that can characterize student's knowledge entirety were proposed and tested; new complex method to visualize the research results is described. Some pedagogical regularity was discovered during the analysis of digesting the course's foundation.

General Plan of Concepts' System Research

In order to examine the entirety of concepts, formed after studying the course, the following strategy was developed and realized.

1. Using existing textbooks (see books [Tanenbaum, 1998; Hamacher et al., 2001; Cilker et al., 2004 and Broydo et al., 2006] for example) and personal teaching experience [Eremin, 2003], the author formed a list of basic terms and concepts. It was assumed, that the rate of digesting and mastery of these terms verified the success of the course.

2. Relying on these resources, the most significant links between terms from the list were fixed. This part of work had an evaluative character and was aimed at a theoretical goal a scrupulous student should achieve. The results of this analysis and course's contents systematization were published in [Eremin, 2007 and Eremin, 2008].

3. Special computer program was written to check the associations between the concepts that students have and fix them in the text file, suitable for further computer analysis. This program uses as input data the list of terms, which was obtained after implementation of item 1, and the list of relations, worked out during the analysis in item 2.

4. Another computer program allows experimenting teacher to analyze the results of knowledge checking, get in the previous item, and output different statistics. Further this software was modified for drawing graphic representation of calculated data.

Let us review these stages of research more specifically.

Organization of Experiment

The full list of concepts for the architecture course, used as experimental base, contains more than 120 terms. The most general concepts – like *computer*, *software* and *hardware*, *theoretical foundations* etc., complemented by the terms that expand the previous ones – *operating system*, *processor*, *memory*, *DMA*, *principle of hierarchy*, *byte* and many others, were included in this base. The terms from related disciplines such as microelectronics, logics and number notations were added into the list. Contrary, the list did not include the names of concrete operating systems, external devices and their manufacturers, and other similar data, less essential from the position of learning the main course's regularities. Using the standard terminology from object-oriented programming, we may say, that classes of the concepts were under consideration, but not their instances.

The wide list of concepts, obtained as described above, forms the set of terms that, from the lecturer's point of view, a literate student must know and understand. The list was found to be large enough, so, as further experiments shows, competent students usually used a little more than the half of it in their answers.

The next step was to analyze **the links between the selected terms**. Non-trivial result of this work consists in the fact, that a limited set of relations was enough to describe the subject. The final set consists of standard relations such as, for example, *whole/part* or *class/subclass* as well as several links, specific for the course, like *base* (*principles of hierarhy* and *addressing – base – memory*, *program counter – base – the main algorithm* of instruction execution) or *connection* (*controller – connection – bus*, *input devices – connection – system block*). The full table of relations with concrete examples for each can be found in publications [Eremin, 2007 and Eremin, 2008]: it consists of only 11 base associations.

The results of analysis, published earlier, are of independent interest. From the point of view of this paper, constructed lists are original data, which are used during the experimental testing of student mastery of the

knowledge domain under consideration. As knowledge control was realized with the help of computer, both lists (of concepts and relations), being input data for control program, were saved in the form of text files.

Now let us proceed to the description of **computer software for students' knowledge check**. It was developed in Delphi programming environment and is simple enough. There are three lists (see fig. 1), using which a student forms a relation looking the following way:

term 1 – relation – term 2

(for instance, the relationship *processor* – *whole/part* – *register* can be easily decoded as phrase «processor and register are linked by association whole/part» or, more exactly, «register is a part of processor»). After student constructs a linked pair of concepts, s/he fixes it clicking the control button «Fix this link». Then the constructed text is added to multi-line text field, arranged in the bottom of the program window.

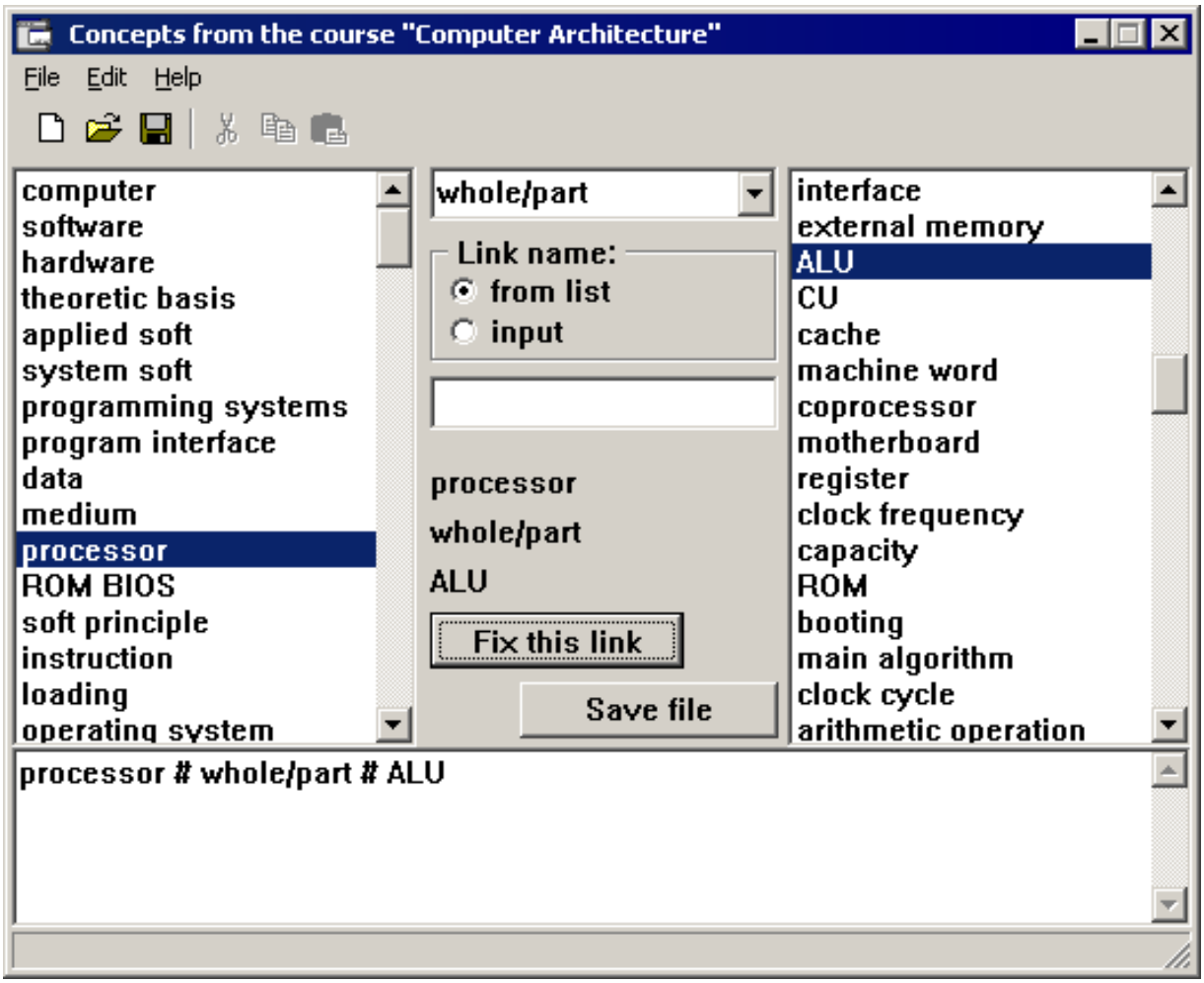


Figure 1

Except the list of associations, loaded from pre-arranged file, the program covers a potential possibility to input additional (not provided by the author) name of relation between the terms: special radio buttons and text field in the center of the window allow such input. As experiments shows, students did not use this alternative, preferring to select link type from the available list.

When a student finishes his/her work, the results are saved in a text file, designed for the further computer analysis. Such testing of knowledge was realized twice – before learning the course and after it: the research was aimed at comparison of these results for assessment of the course digesting.

At the last step all text files with the results of knowledge control fall under **computer processing by means of the second program**. Its main aim was to educe linked groups of concepts for every student. For example, when computer, processing a file, finds associations *functional units – class/subclass – processor*, *functional units – class/subclass – memory*, *functional units – class/subclass – input devices* and *functional units – class/subclass – output devices*, it joins all five concepts, mentioned there, into one group. Later on some other terms will be added to this group: for instance, relations *processor – whole/part – ALU* and *processor – whole/part – CU* join two new terms to the group – *arithmetic and logic unit* and *control unit*.

In ideal case all concepts of the course are interrelated; test running of the program with thoroughly prepared author's file, built according to full results from publications [Eremin, 2007 and Eremin, 2008], confirms this. But experiments show, that real students' files represented more scattered picture, which consists of several isolated groups of concepts, and some groups were very small (2-3 terms). Such small groups must be interpreted as a separate fact that student does not associate with other facts from the course. You must note that an increased rate of fragmentation indicates student's knowledge is sparser.

Experimental check of knowledge entirety was organized the following way.

The experiments estimated knowledge of students, who studied on the third course of the physical faculty. Knowledge control took place twice: at the beginning and at the end of semester, that is to say before and after the learning of the referenced discipline «Computer Architecture». Unfortunately, the number of students, who learned the course and took part in the experiment, was small. Additionally, several students missed one of the tests for a variety of reasons, so their results were incomplete and could not be considered.

Students did not know that the aim of the experiment was entirety of concepts' system because it could artificially improve their results. They were simply told that testing tries to verify the rate of digesting the material. Their instruction was not to think about the parameters of evaluation, but just try to demonstrate digested knowledge the best way they could.

As the students did not study the types of relations between concepts before, they were given a special table (with all 11 referenced links and numerous examples for each one) before testing. The results shows, that it was insufficient and the students unsatisfactorily differentiated the types of links. Often they even missed classical relations *whole/part* and *class/subclass*, not to mention other link types. The aim of the experiments was to

estimate the general entirety of the basic concepts – so from this point of view concrete kinds of links are not too important. Such effect became a motive to neglect the errors in this part of the task on the first stage of experiments, and just fix the existence of relation but not it type. This simplification of method notably facilitated the process of results processing and analysis.

The time for implementation of the task was not limited, the work finished individually at students' will. Presumably, it brought some uncertainty into experiment, since some students really indicated all interrelations they knew, but others just got tired and finished their work. The task execution employed about an hour in average. According to my observation of computer testing, the students' reaction was mainly neutral («we just get one more task»), so the checking procedure did not lead to any difficulty.

Discussion of Results

As it follows from described experimental method, the results are fixed in a text file with pairs of concepts, linked together with association of some kind. It was also validated above, that the links' types are not considered on this stage of research. Special software was developed for processing of final files; it combined related concepts into groups and then calculated a set of statistical characteristics.

Let us discuss what parameters can claim to be the characteristics of student's knowledge entirety.

Primary parameters are evident: *total number of terms* and *total number of links* between them; they can be easily counted from any student's file with the results. The ratio of these values, that means *average number of links per one concept*, also may be introduced into consideration. It is evident, that the more these values are, the better student mastered the material.

Another set of parameters may be built while arranging interrelated terms into groups. We can offer here *total number of concepts' groups* (this value must be as small as possible, in ideal case all terms must form the only group) and *size of the largest group* (this factor we want to see as large as possible). Additionally we can divide total number of terms by number of groups, i.e. get *average size of group*, which must be large when student learned the course profoundly.

All listed above characteristics were calculated for every student, and then compared for testing before and after the course. The most interesting seems to be the results for **average number of terms per group** (T/G). Unfortunately, limited number (N=12) of students, who took part in the experiment, does not allow to validate with statistical significance, that this factor is the best integral parameter for the process.

Diagram of values for selected measure is presented on fig. 2. Let us consider this picture more detailed.

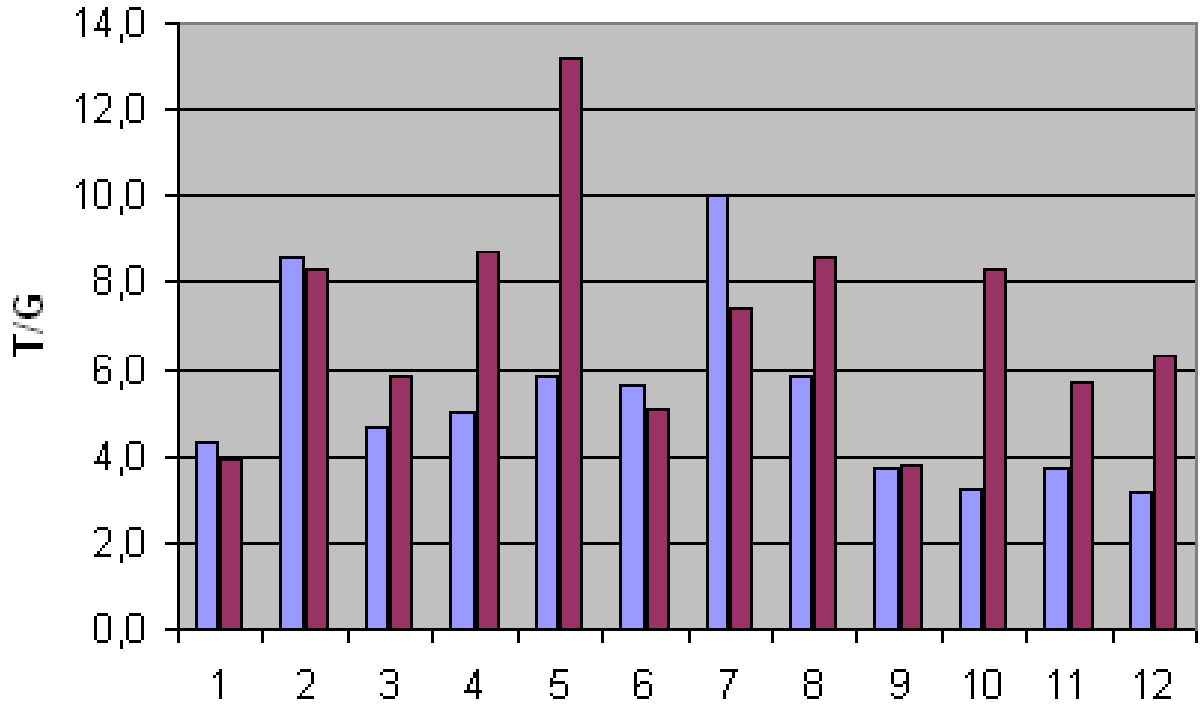


Figure 2

Fig. 2 shows the results of knowledge entirety check for 12 students (points on the X-axis are marked by their numbers from 1 to 12). In Y direction diagram's columns are built: their height are proportional to average size of concepts' groups, i.e. total number of terms T divided by number of groups G, formed from them. The left columns describe data before learning the course, and the right ones – after its studying.

It's important to mention that students on the diagram are rank-ordered according to some rating: the criterion of such arrangement was time of finishing all the tasks, given by the teacher. The students with small numbers finished the course earlier; hence they are supposed to demonstrate better results in learning the course content. In opposite, columns for «the slowest» students form the right part of the picture. Subjective impressions and interview with students during learning process confirm the acceptability of selected criterion as measure of learning success, at least for this group of students.

Let us analyze the values of average concepts in a group, to say more exactly, their changing after finishing the course. It is clearly seen from the diagram, that students have different changes of factor T/G. Some of them (for instance, student number 1, 2, 6, 9) have practically the same values before and after learning the course, but some others (4, 5 and 10-12) show notable growth. It is interesting to mention one feature of the diagram: all weaker students, who have large numbers, improved the entirety of their knowledge, but students with the best rating, contrary, did not demonstrate significant growth. The last fact may be interpreted as good preliminary knowledge of such students – they had some experience in computer architecture before the beginning of study. Unfortunately, the available data are not sufficient to make a pedagogic generalization.

We must pay attention to the magnitude of average group size. As fig. 2 shows, the best result for investigated students lies near value 13, whereas in ideal case for set of terms, used in experiment (more than 100), it can not be less than the last value.

One more (fundamentally new) form of visual representation for entirety of concepts' system for the same group of students is demonstrated on fig. 3.

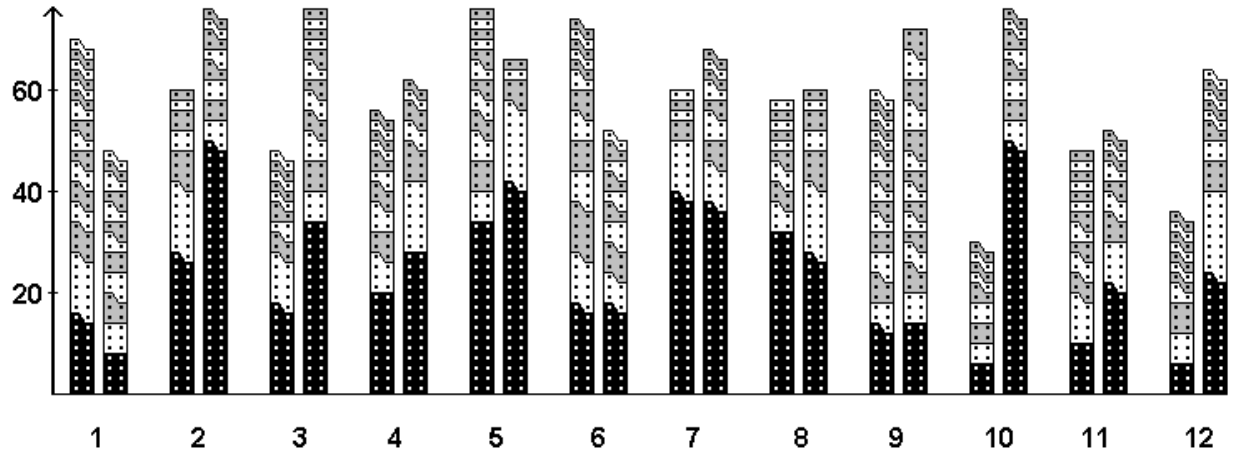


Figure 3

This «spotted» diagram is organized the following way. Pairs of columns we see on it represent input and output testing: the right bar in every pair indicates final results. Height of any column is proportional to the number of concepts that student selected during assessment. Every dot in a column means one individual concept. All columns are divided into several areas; each of them represents a group of interrelated concepts, learnt by every student. For better visibility, neighbor areas are painted in white and gray colors. The black region in the bottom, which is always the largest, symbolizes the kernel of student's knowledge. As you can notice, all groups in every bar are regularized by size, so the smallest groups from 2-3 concepts (such groups may be interpreted as separate facts out of common picture) are always placed to the top of the bar.

Thus form of complex visual data representation, proposed in this work, clearly reflects the following information:

- the height of the columns is proportional to the number of the learnt concepts, which have at least one relation with the other terms;
- the number of the multicolored areas characterizes the rate of grasped data's scattering (fragmentation indicator) in selected knowledge domain;
- the size of interrelated concepts groups (proportional to their square) also shows the entirety of student's knowledge in analyzing part of the course;
- size of the largest area in the bottom of the column (filled with black color on fig. 3) characterizes the volume of the basic block of course's concepts.

As it was mentioned above, in ideal case the diagram bar must be heterogeneous black bar (consisting from the only group), and its height must include all the concepts of the course. Real picture, as you can see from fig. 3, is far from ideal: students' knowledge comes apart on several independent groups of terms and the highest bar includes less than 70 concepts from more than 120 ones that were offered in the task.

Examining fig. 3, we may get several practical conclusions about successfully digesting of the course content. The results for students number 4-9 and 11 (numeration is the same as on fig. 2) had slightly changed after learning the course; students 4, 5 *и* 11 stand out of the list, because the structure of their diagrams is rather better: the number of areas decreases, and their width, in opposite, increases. At the same time students 2, 3, 10, 12 demonstrate more essential growth of factors. And one more observation. The height of some bars (especially for students 1 and 6), which depends from the number of digested concepts, fall off after the course. We may assume, that this paradox can be explained by students' intention to pass the last test quicker and get their long-expected credit. Of course, this assumption needs in further experimental check.

At last, carefully analyses shows, that there are some identical columns on fig.3: for instance, compare right columns for students 2 and 10. This means that one of them cheated: just copied the result file from other student. Unfortunately, I noticed it too late to enforce students to rewrite their work. So anti-cheating measures must be developed for further experiments.

Conclusion

The computer method for experimental research is offered, that allows to study the entirety of system of basic concepts after educational course learning. This method was tested on students' learning the course content in «Computer Architecture» and the results showed its efficiency. This paper also describes a new original visual form of data representation, which clearly demonstrates the structure of interrelations between concepts in student's knowledge. Some interesting preliminary pedagogical results were discovered during the research, for instance, how knowledge entirety depends on the level of student's background. Work on the research method and a more detailed study of the process of how basic concepts' system is formed will be continued.

Acknowledgement

The paper is partially financed by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria. www.ithea.org, www.foibg.com.

The author expresses his thanks to J. Caruso for helpful discussion of the research results.

Bibliography

[Broydo et al., 2006] V.P.Broydo, O.P.Ilyina. Architecture of computers and systems. Piter, St-Petersburg, 2006 (in Russian).

[Cilker et al., 2004] B.Ya.Cilker, S.A.Orlov. Organization of computers and systems. Piter, St-Petersburg, 2004 (in Russian).

[Eremin, 2003] E.A.Eremin. Popular lectures about computer organization. BHV-Petersburg, St-Petersburg, 2003 (in Russian).

[Eremin, 2007] E.A.Eremin. Using Topic Map technology in the planning of courses from the CS knowledge domain. In: Proc. Seventh Baltic Sea Conference on Computing Education Research (Koli Calling 2007). CRPIT, 88. ACS, 2007.

[Eremin, 2008] E.A.Eremin. The analysis of content line "Computer" of the informatics course with application of computer methods of knowledge representation. In: Informatics ("1 September"), N 9, 2008 (in Russian).

[Hamacher et al., 2001] C.Hamacher, Z.Vranesic, S.Zaky. Computer organization. McGraw-Hill, New York, 2001.

[Kushnirenko, 1998] A.G.Kushnirenko. Several remarks about school informatics in Russia of 1998 year. In: Informatics ("1 September"), N 16, 1998 (in Russian).

[Podlasiy, 2002] I.P.Podlasiy. Pedagogy: the new course. Book 1: General foundations. Learning process. Humanit. ed. center VLADOS, Moscow, 2002 (in Russian).

[Tanenbaum, 1998] A.S.Tanenbaum. Structured computer organization. Prentice Hall, Upper Saddle River, NJ, 1998.

[Zorina, 1992] L.Ya.Zorina. Textbooks on sciences foundations. In: What a Textbook Must Be: Didactic Principles of Composition. P.2. Moscow, 1992 (in Russian).

Author's Information

Evgeny A. Eremin *– Senior lecturer, Perm State Pedagogical University, 614990, Sibirskaya st., 24, Perm, Russia; e-mail: eremin@pspu.ac.ru*

Major Fields of Scientific Research: Knowledge representation, Computer Learning, Computer architecture

ZOOMING USER INTERFACE IN PRESENTATIONS FOR LEARNING

Artem Pecheny

Abstract: *The paper describes interactive environment for creating and using in learning multimedia presentations of compound visual structure with linking arrows due to zooming user interface. Therefore the environment can be viewed as hypermedia one. The article contains the analyses of appliance well established principles of modern instructional theories as well as hypotheses in application hypermedia in learning.*

Keywords: *cognitive load theory, cognitive theory of multimedia learning, zooming user interface, hypermedia.*

ACM Classification Keywords: *K.3.1 Computer Uses in Education*

Introduction

Many studies describe the usage of presentation in learning. This part of using Information and communications technology (ICT) is important because every teacher is able to create presentations. The main features of presentation are usually associated with MS PowerPoint capabilities. Theoretical findings also have appliance in practical guidelines for traditional presentation creation [Atkinson & Mayer, 2004]. Taking into account theoretical principles we try to find a different approach to the subject. We use Zooming user interface as the main idea of our system and support this idea by set of tools following the findings of theories. Students of Perm State Pedagogical University (Russia) make presentations in our system in the topic of physics. These presentations can be used in different activities such as lectures, individual studies or team work. We assume that interactive whiteboard should be used in a classroom environment for the best result.

Theoretical foundation

There are few of the most influential theories in the area of instructional design and computer-based training: The Cognitive load theory (CLT) [Sweller et al., 1998] and The Cognitive Theory of Multimedia Learning (CTML), [Mayer et al., 2000].

Cognitive load theory is mainly concerned with the learning of complex cognitive tasks, where learners are often overwhelmed by the number of information elements and their interactions that need to be processed simultaneously before meaningful learning can commence [Paas et al., 2004]. From CLT point of view the

learning can be considered as schema construction and automation [van Merriënboer and John Sweller, 2005]. Schema is one of the concepts for representing mental structures.

Sweller et al. [1998] proposed several instructional design techniques based on Cognitive Load Theory. These instructional principles are identified as the *goal-free effect, worked example effect, completion problem effect, split-attention effect, modality effects, redundancy effect,* and *the variability effect*. All of these effects except *variability effect* are meant to reduce extraneous cognitive load, which is not necessary for learning but demands mind resources.

We give description of the effects that can be implemented in our environment following van Merriënboer and John Sweller [2005] :

Goal-free effect: Replace conventional problems with goal-free problems that provide learners with an a-specific goal.

Worked example effect: Replace conventional problems with worked examples that must be carefully studied.

Completion problem effect: Replace conventional problems with completion problems, providing a partial solution that must be completed by the learners.

Split attention effect: Replace multiple sources of information (frequently pictures and accompanying text) with a single, integrated source of information.

Modality effect: Replace a written explanatory text and another source of visual information such as a diagram (unimodal) with a spoken explanatory text and a visual source of information (multimodal).

Redundancy effect: Replace multiple sources of information that are self-contained (i.e., they can be understood on their own) with one source of information.

One more effect was found later:

The expertise reversal effect: Instructional methods that work well for novice learners may have neutral or even negative effects when expertise increases.

The Cognitive Theory of Multimedia Learning [Mayer, 2001] which is partially derived from Dual Coding Theory [Paivio, 1978] draws a quite conclusive picture concerning learning with visualizations. The theory assumes that there are two ways of processing information and hence two kinds of mental representation in the cognitive system. In the verbal system, information of a sequential structure such as written texts or spoken words is processed. In the non-verbal system, spatial information and pictures are processed. Connecting these two cognitive representations properly should improve learning results. [Hoffler, 2010].

The cognitive theory of multimedia learning applies some basic assumptions to the design of multimedia learning environments that are similar to CLT. In addition, this theory defines several types of sequential processes that are required for active learning [Kalyuga, 2011].

Applications of this theory are theory-based principles for how to design electronic learning environments, which themselves can be tested in research studies. As part of his evidence-seeking efforts for the science of e-learning, Mayer [2001, 2003] presents nine major effects which were developed out of dozens of studies. These replicated effects are: *modality effect, contiguity effect, multimedia effect, personalization effect, coherence effect, redundancy effect, pre-training effect, signaling effect,* and *the pacing effect*. An explanation of part of these effects concerned with using the environment, lead to principles [Moreno & Mayer, 2000], follows:

The Contiguity Principle states that better transfer occurs when corresponding narration and animation are presented simultaneously, both temporally and spatially. Temporal contiguity means that corresponding words and pictures should be presented at the same time, while spatial contiguity means that corresponding words and pictures should be presented near rather than far from each other on a page or screen.

The multimedia principle states that better transfer occurs from animation/pictures and narration/words than from words alone. When words and pictures are both presented, learners have the chance to construct verbal and visual cognitive representations and integrate them.

The redundancy principle states that better transfer occurs when animation and narration are not combined with printed text. When pictures and words are both presented visually, it can overload visual working memory capacity.

The pacing principle states that better transfer occurs when the pace of presentation is controlled by the learner, rather than by the program. Learners vary in the time needed to engage in the cognitive processes of selecting, organizing, and integrating incoming information, so they must have the ability to work at their own pace to slow or pause the presentation if necessary. If the pace of the presented material is too fast, then these cognitive processes may not be properly carried out and learning will suffer.

These principles are general enough and can be well applied in particle designs. But we try to develop educational environment which enable implementation of these principles more natural way. On the other hand we try to use findings from different less established fields. One of them is Hypermedia environment research, which is more arguable and has pros and cons. One of hypermedia aspects is learner control, which can be also associated with form of interactivity and can be discussed from CTML point of view. Scheiter and Gerjets [2007] note that learner control in hypermedia context is supposed to aid learning, because respective environments (1) mirror the human mind, (2) increase interest and motivation to learn, (3) enable instruction adapted to learners' preferences and cognitive needs, (4) provide affordances for active and constructive information processing, and (5) foster the acquisition of self-regulatory skills.

It is also mentioned in this work that the potential effectiveness of self-controlled learning with hypermedia is hampered by (1) usability problems (i.e., disorientation, distraction, cognitive overload), (2) moderating learner characteristics (i.e., prior knowledge and general abilities, self-regulatory skills, cognitive styles and attitudes

towards learning), (3) a lack of conceptual foundations, and (4) methodological shortcomings of hypermedia research.

Usability problem is a separate area and its description is out of bounds of this article, but some aids will be presented further. Some studies of learner characteristics and cognitive styles are of interest. One of these studies investigated the effect of two segmented and holistic, on the progression over time of learners' mental models toward that of an expert with the moderator of cognitive flexibility [Aubteen Darabi et al. 2009]. Authors suggest that efficient choice of segmented and holistic strategy is closely concerned with cognitive flexibility, one of cognitive styles characteristics. This conclusion is consistent with Swanson and Law's [1993] whole-part-whole instructional model. That assumes through "first whole" the model introduces new content to learners by formatting in their minds the organizational framework, then "parts" learned separately, and after learners have successfully achieved the performance criteria for individual "parts" within the whole, the instructor links these parts together, thus forming the second whole.

Zooming user interface

A ZUI is a dynamic interface. It provides user a canvas that is larger than the viewing area, on which items are placed: The items placement is arbitrary and may be determined by the users, by the system, or by both. User can scroll their viewing window across this canvas to view different items. The ZUI difference from a normal canvas is that the user may zoom in or out onto a particular item, like a telescopic lens on a camera [Salmoni, 2004].

In application of presentation construction ZUI means that there are no separate slides more. They turn to frames – dedicated parts of large canvas that can be placed arbitrary, can have arbitrary scale. Smaller frames can be combined in a larger one. And all of them can be linked by pointing arrows (Fig. 1, 2, 3). Thus, all of presentation content can looks like concept map and represent connections between parts. This approach assumes, that titles of frames are large enough to be readable in low scale, when presentation or its part presents structure of the content only.

Zoom environment has good consideration with the Whole-part-whole learning model: the "first whole" is introduced in low amount of scale, then "parts" are surveyed closely, and then each "part" finds it's place in the "whole" again.

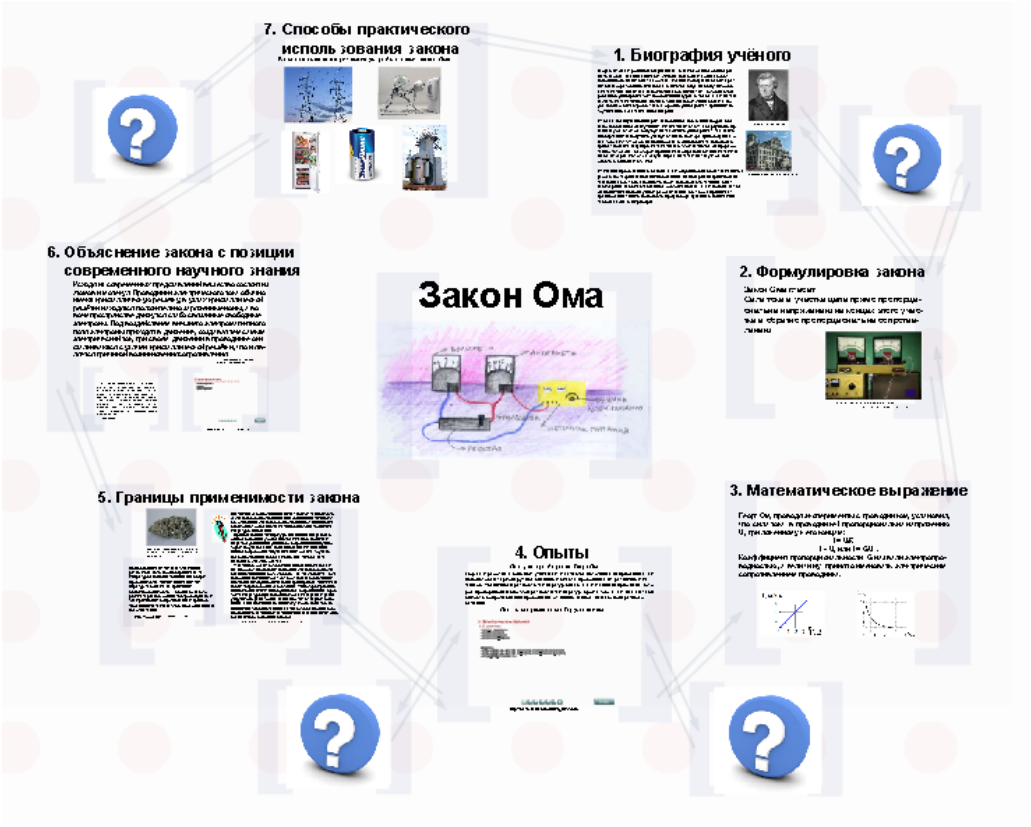


Fig.1. View of Zooming Presentation about Ohm's Law by Mila Khrenova, student of PSPU.

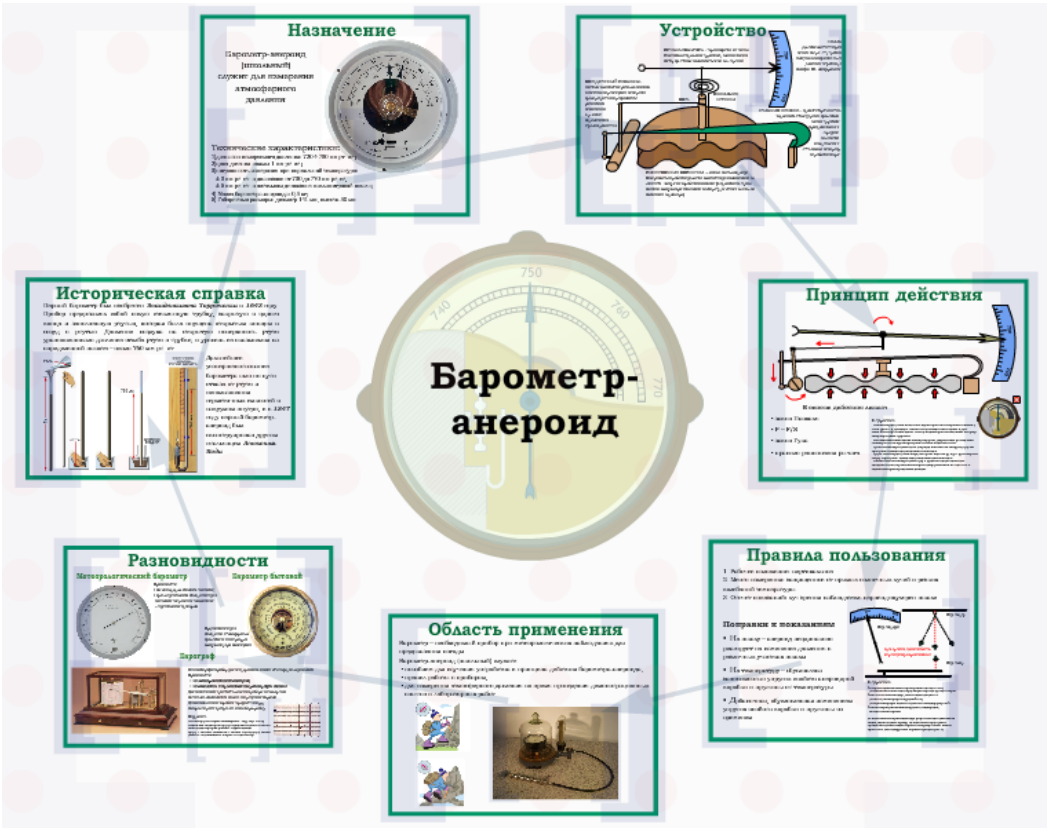


Fig.2. View of Zooming Presentation about aneroid barometer by Stas Subbotin, student of PSPU

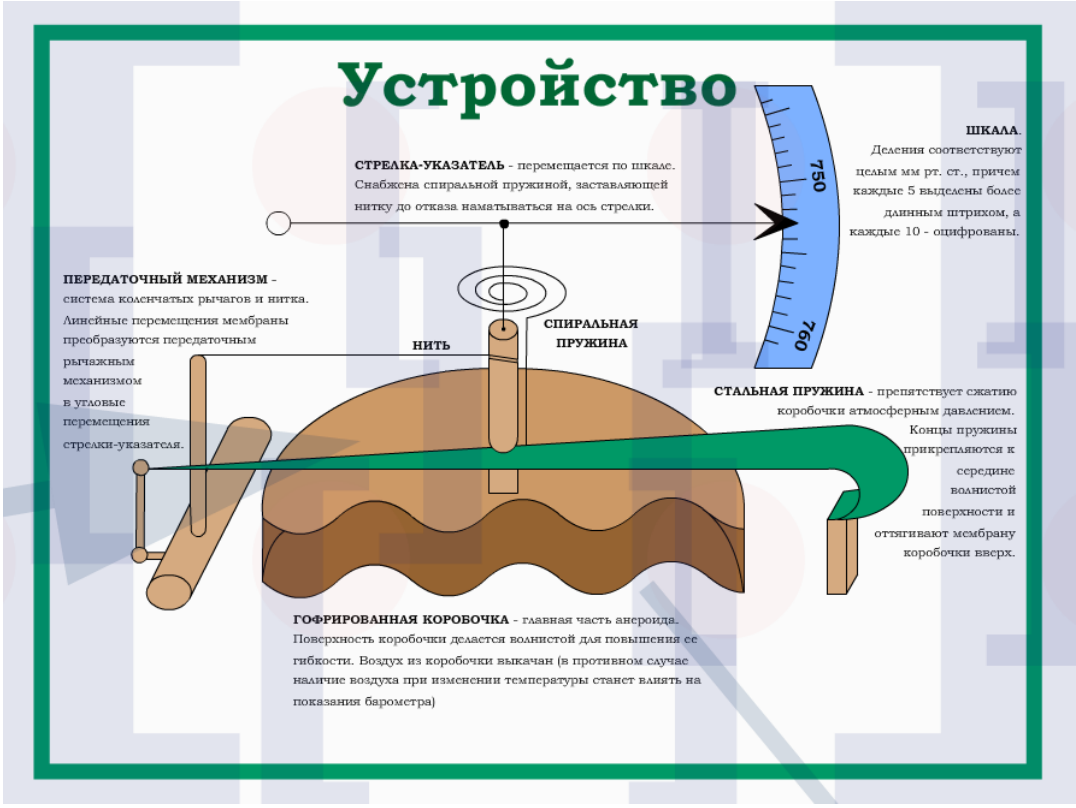


Fig.3. Zoomed fragment of Presentation about aneroid barometer

Features of the environment

First of all we should note that elements of described presentation can be in almost any media format: text, pictures, video, animations, sounds and anything that can be placed in SWF format (interactive models, 3D models, units displaying images from WEB-camera, etc). Arrows in presentation also take function of links – clicking on arrow will place pointed frame in viewing area. Thus, our environment match description of Hypermedia learning environments: "…consist of network-like information structures, where fragments of information are stored in nodes that are interconnected and can be accessed by electronic hyperlinks" [Scheiter and Gerjets , Conklin 1987; Rouet et al. 1996].

Now let us describe set of tools realized in the environment. Main from them are: *painting tool, mask layer tool* and *drag-and-drop*.

Painting tool allow presentation's author to set "drawable" property to any element of presentation (f.e. picture or movie). Then, when presentation is being viewed, user can draw marks on this element by mouse pointer just like in a graphics editor without any additional actions. In case of pictures or animations it can be explanatory statements, or, in case of clean fields it can look like a part of simple whiteboard, which allows learner to write task solution.

Any parts of presentation can be overlapped and one can hide another one. If *mask layer tool* is assigned to top element, it is possible to cut holes in this element to make the bottom element visible.

And *drag-and-drop tool* means that element of presentation can turn to "dragable" element, so a number of those elements can be reordered as a part of some learning task.

Implications of theoretical foundation

In this section we show appliance of multimedia principles in creation of multimedia presentations in our environment.

Goal-free exploring strategy in hypermedia environment is very natural, so it is closely concerned with *goal-free effect*. It is an efficient way to introduce learners in complex topic.

Besides usual way *worked example effect* as well *as completion problem effect* can be freshened up with painting tool, layered masking tool and its combination. Any part of a solution can be hidden, any part can be dedicated to be filled by learners. According to *split-attention effect* and *contiguity effect*, zoom environment allows to place any elements to single construction. Even extensive description can be integrated with images due to zooming placement. Following *redundancy principle* we should exclude printed text which is presented in narration, but in some cases due to *expertise reversal effect*, printed text can be preferable. In zoom environment that redundant texts can have low scale, so it will be unnoticeable usual way, but can be accessible on request. Also some cases *mask layer tool* can be applied well. *Modality effect* and *multimedia effect* take their place when images, video, animations and sound records are used. *The pacing effect* appears firstly due to leaner's control in hypermedia environment and secondly due to splitting sound records by elements they targeting. Usability problems (i.e., disorientation, distraction, cognitive overload) are partly resolved due to possibility to take zoom out any time and to view all system of topic knowledge and to find placement of current element.

Conclusion

The paper presents the description of interactive environment for creating and using in learning multimedia presentations. It shows how multimedia principles can be applied for creating presentations in the environment. The assessment of student's work will be discussed in a separated paper.

Bibliography

[Atkinson & Mayer, 2004] Atkinson, C., & Mayer, R. E.. Five ways to reduce PowerPoint overload. Creative Commons, 1-15, 2004. Retrieved from http://nursing412.pbworks.com/f/atkinson_mayer_powerpoint_4_23_04.pdf

[Aubteen Darabi et al., 2009] Aubteen Darabi, Jennifer Hemphill, David W. Nelson, Wilma Boulware, Xinya Liang. Mental model progression in learning the electron transport chain: effects of instructional strategies and cognitive flexibility. Adv in Health Sci Educ, 2009. DOI 10.1007/s10459-009-9212-0

[Conklin, 1987] Conklin, J. Hypertext: An introduction and survey. IEEE Computer, 20, 17–41, 1987.

[Hoffler, 2010] Tim N. Hoffler. Spatial Ability: Its Influence on Learning with Visualizations—a Meta-Analytic Review. Educ Psychol Rev (2010) 22:245–269, 2010.

[Kalyuga, 2011] Kalyuga, S. Cognitive Load Theory: How Many Types of Load Does It Really Need? Educ Psychol Rev. DOI: 10.1007/s10648-010-9150-7, 2011.

[Mayer, 2001] Mayer, R. E. (). Multi-media learning. New York: Cambridge University Press, 2001.

[Mayer, 2003] Mayer, R. E. Elements of a science of e-learning. Journal of Educational Computing Research, 29(3), 297-313, 2003.

[Merriënboer and Sweller, 2005] Jeroen J. G. van Merriënboer and John Sweller. Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. EDUCATIONAL PSYCHOLOGY REVIEW. Volume 17, Number 2, 147-177, 2005.

[Moreno & Mayer, 2000] Moreno, R., & Mayer, R. E. A learner-centered approach to multimedia explanations: Deriving instructional design principles from cognitive theory. Interactive Multimedia Electronic Journal of Computer Enhanced Learning, 2000. Retrieved March 22, 2005, from http://imej.wfu.edu/articles/2000/2/05/index.asp

[Moreno & Mayer, 2003] Mayer, R. E., & Moreno, R.. Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist, 38(1), 43-52, 2000.

[Paas et al., 2004] Paas, F., Renkl, A., Sweller, J. Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. Instructional Science. Vol. 32. Springer Netherlands, 2004

[Paivio, 1986] Paivio, A. Mental representations: A dual coding approach. Oxford, England: Oxford University Press, 1986.

[Rouet et al. 1996] Rouet, J.-F., Levonen, J. J., Dillon, A., & Spiro, R. J. Hypertext and cognition. Mahwah, NJ:

[Salmoni, 2004] Alan James Salmoni. The Zooming User Interface, 2004. Retrieved 2010 from / http://www.advogato.org/article/788.html

[Scheiter & Gerjets, 2007] Scheiter, K. & Gerjets, P. Learner Control in Hypermedia Environments. Educ Psychol Rev, 2007 19:285–307

[Sorden, 2005] Stephen D. Sorden. A Cognitive Approach to Instructional Design for Multimedia Learning. Informing Science Journal Volume 8, 2005

[Swanson & Law, 1993] Swanson, R. A., & Law, B. Whole-part-whole learning model. Performance Improvement, 6(1), 43–53, 1993.

[Sweller et al. 1998] Sweller, J., van Merrienboer, J. J. G., and Paas, F. Cognitive architecture and instructional design. Educ. Psychol. Rev. 10: 251–296, 1998. Erlbaum, 1996.

Authors' Information

Artem Pecheny – PhD Student, Perm State Pedagogical University, Russia;

email: pecheny@gmail.com

Major Fields of Scientific Research: cognitive science, computer graphics

CONCEPTUAL KNOWLEDGE MODELING ON THE BASIS OF NATURAL CLASSIFICATION

Mikhail Bondarenko, Kateryna Solovyova, Andrey Danilov

Abstract: It is difficult to exaggerate the importance, the urgency and complexity of "good" classifications creation, especially in knowledge management, artificial intelligence, decision making. To what extend it is possible within a short paper, the peculiarities and advantages of the new system method of the systemological classification analysis for the classifications of concepts creation were discussed. It is noted that the applying of the natural classification criteria improves considerably the quality and the power of the classification knowledge models and ontologies, allows taking into account the deep knowledge of any, including ill-structured, domains. In the process of the research conduction the system models of the domain fragment of the ontologies on the basis of the parametric classification were created. Some results of the actual domain "Social Networks in Internet" analysis and modelling and the ontology fragments, realized in the ontologies engineering tool Protégé 3.2, are also considered. The systemological classification analysis application has allowed proving the obtained classifications of social networks functions, taking into account the objects essential properties. It has also successfully recommended itself for deep knowledge acquisition; the basic hierarchy of classes, "good" classifications and ontologies creation; possesses predictive power, simple logically relevant structure, ensures the possibility of the correct inference on knowledge.

Keywords*: conceptual knowledge, knowledge systematization, natural classification, ontology, systemological classification analysis, social network, hierarchy, systemology, artificial intelligence.*

ACM Classification Keywords*: 1.2 Artificial Intelligence – 1.2.6 Learning: Knowledge Acquisition*

Introduction

The development of knowledge management, artificial intelligence, decision making and many other actual scientific and practical directions is determined by knowledge and its quality. As we know, knowledge, intellectual capital is the main competitive advantage, the foundation of modern organizations, enterprises, society, human and nations' welfare and important component of decision making support systems.

In different spheres of knowledge acquisition and application conceptual models of subject domains play a leading role. "Historically," the species of domain models are: dictionaries, thesauri (in linguistics), conceptual models (infological, semantic models - in databases), UML diagrams (of classes, of use cases, ... - in object-

oriented analysis and modeling), models of knowledge (semantic nets, frames, ... - in artificial intelligence), ontologies (from the viewpoint of the realization and application one of the most modern kind of a domain model, aimed primarily at the knowledge application in Internet).

The basis of such models is the relationships of the hierarchy between concepts (concepts classification), in the first place, the relations *genus-species* and *part-whole*, about two millennia known in formal logic. These relations in the theory of classification are called the relations of *taxonomy* and *meronomy*, in artificial intelligence – *genus-species: Isa* (class - class), *Instance-of* (class - element) and *part-whole: Part-of*; in object-oriented analysis and modeling – *generalization / specialization* and aggregation (in some cases, *composition*), respectively, etc. In systemology to these relations corresponds one *relation of the functional ability of the whole support*, respectively, for system-classes and concrete systems (which are reflected in general and single concepts).

How effective are the methods of the concepts classification creation - the basis of modern models of knowledge of domains? The analysis shows that in most domains the classifications are subjective; many of them do not meet even the requirements of formal logic. That is why it is proposed to apply a new unique method of the systemological classification analysis based on the natural classification [E. A. Solovyova, 1999; E. A. Solovyova, 1991; E. A. Solovyova, 2000], which has successfully recommended itself for deep knowledge acquisition, the basic hierarchy of classes, "good" classifications and ontologies creation in all, including ill-structured domains.

Introduction to the Natural Classification as the Conceptual Knowledge Systematics

As noted, this work is not about data classification into existing classes. We work with deep knowledge classifications and besides with the conceptual deep knowledge, on the conceptual level, determine classes (entities), properties and relations, and besides in accordance with their position in the domain, in the reality, in accordance with the systemic of the reality. Naturalists and other scientists interested for many centuries in the problem of "good" classification creation, *the position of objects in which reflects the reality (the domain), is determined by essential properties and relations of objects* and therefore possessing predictive power. This "good" classification was called systematics, or the natural classification, the first meaningful criteria of which were introduced by the Englishman Wavell more than 150 years ago; then by A. A. Liubishchev, Y. A. Schrayder and other scientists, for example, the natural classification - *is a form of the laws of nature presentation..., expresses the law of the systems of reality relationship, allows to reach the maximum number of goals, because it takes into account the essential properties, etc.* Such criteria are useful for fundamental science, but are not constructive for computer modeling, application in knowledge models and ontologies. That is why in Knowledge Acquisition Laboratory, at the Social Informatics Department and Scientific-Educational Knowledge Management Center for more than 20 years the systemic research of conceptual knowledge and natural classification has been conducted. For the first time the constructive criteria of the natural classification and a new method of systemological classification analysis which allow to take into account deep knowledge, objects essential

properties and relations in domain models in the most objective way, have been obtained [E. A. Solovyova, 1999; E. A. Solovyova, 1991; E. A. Solovyova et al; 2000, E. A. Solovyova, 2000, etc.]. This method for the first time synthesizes system and classification analysis. The natural classification criteria correspond completely to the formal-logical criteria and also deepen and generalize them.

These fundamental results have not only theoretical but also an important practical value. They allow creating knowledge models and ontologies which take into account essential properties and causal-investigative relations, possess predictive power, simple logically relevant structure, allow generalization and unlimited knowledge refinement without redesigning classification, ensure the possibility of the correct inference on knowledge, recommendations and decisions making support, interface with the concepts of natural language application.

It is proved mathematically and systemologically and (with the use of the category theory and the categorical-functorial model of the natural classification obtaining) that the natural classification is the parametric one (including properties of all its elements), in which the properties classification determines (isomorphic) the objects classification, the properties properties classification – deep layer poperties – the properties classification, etc.). In practice, the consideration of one level of properties (their genus - species classification) allows making the classification model founded and really effective for solving on its basis the various tasks that require knowledge application.

Functional systemology - the systemic approach of the noospheric stage of science development – was created for and is aimed at complex, qualitative, ill-structured problems solving, it differs profitably from the traditional systemic approaches and for the first time really takes into account the systemic effect. Systemology, taking into account the principles of systemic, integrity and diversity, considers all objects, processes and phenomena as systems functioning to support the supersystem functional abilities. Systemology as modern system methodology does not regard system as a set but as a functional object which function is assigned by supersystem. Systemology in particular allows overcoming problems of traditional methods of system analysis at the expense of using conceptual knowledge as well as formalizing procedures of analysis and synthesis of complex systems and creating knowledge-oriented software tools for their simulation. The development of the concrete (internal) systems systemology of G. P. Melnikov for the system of classes allows deep knowledge getting and modeling for all, including ill-structured, domains [Bondarenko et al, 1998; E. A. Solovyova, 1999].

Social Networks Functions Classifications

The conceptual knowledge modelling will be accomplished on the example of the actual domain of **social networks**, including the ontology creation. Nowadays the need to solve complex problems requiring the knowledge of the domain specialists appears increasingly. To train highly qualified professionals progressive companies propose to use the conception of learning organizations. A learning organization as a tool for solving

problems related to the company professional level improving. To create and acquire knowledge the company needs to be constantly in the process of self-improvement. One of the advanced methods of the organization development is the social networks use. The social networks in Internet functions research will allow understanding better the expediency of their use, to use the social networks more effectively in decision making, for further knowledge systematization in the social networks domain.

Resulting from the research the developped social networks classifications were not found. There are several articles where the social networks in Internet functions but not their classifications are mentioned. For example, the following main functions of social networks in Internet are allocated:

- profiles, communities, blogs dogear, activities [Byelenkiy, 2008];

- functions of personal profiles creation, of users interactions, of common goal achieving by means of the cooperation, of resources interaction, needs satisfaction due to the resources accumulation [Kuzmenko, 2009].

The functions research was conducted through the practical use of the network with identifying semantics the functions during the direct work with the network. Due to the weak structuring of the domain and the distinguished functions names in different networks, which in practice often have the same functional destination in most networks, it was necessary to undergo the registration process (or just to come into the network workspace) and to use practically the social network functions to determine their real functional destination. Also during the analysis process, the difficulties have appeared due to not clearly defined functions names, which haw required the additional research and the functions comparative analysis conduction [Danilov, 2010].

The analysis shows that in the first division base, for example, the communication (messages interchange) functions class is absent. In the second division base the search functions are absent and it is also not clear what is meant by the functions of common goal achieving by means of the cooperation. The authors of the given divisions do not exemplify the functions which refer to the classes of these divisions.

Thus, the knowledge systematization in the domain of social networks is needed. Subsequently it will allow not only to obtain the social networks ontology but also to improve the considered nets from the functional viewpoint, to expand the set of their functions, to improve the meaningful placement of the menu functions in concrete social networks. The results of the social networks systematization may be applied for a new social network creation taking into account the advantages and disadvantages of the existing social networks.

In this case we consider the classification creation by the functionality as the knowledge systematization in the given domain [Solovyova, 1999].

Let us consider the comparative analysis of the existing classification functions and the proposed network supposed systematics.

As an example of the hierarcy first level of the social network "В Контакте.ru" functions classification, implemented in the software tool Protégé 3.2. (see Figure 1) is presented.

The created classifcations are implemented in the software tool Protégé 3.2. The choice of this software tool is grounded by the fact that it is a free, open-source ontology editor. Protégé has an open, easy spread architecture at the expense of the functionality extension modules support which are freely available on the Protégé official site. The knowledge model is an OKBC-compatible (Open Knowledge Base Connectivity – it is the application programming interface for the access to knowledge bases), this allows applying in Protégé the one customized interface for different semantic markup languages processing. An example of such language is OWL (Web Ontology Language). All the listed above possibilities of Protégé, as well as the visibility of the obtained classifications was the reason for using namely this software tool.

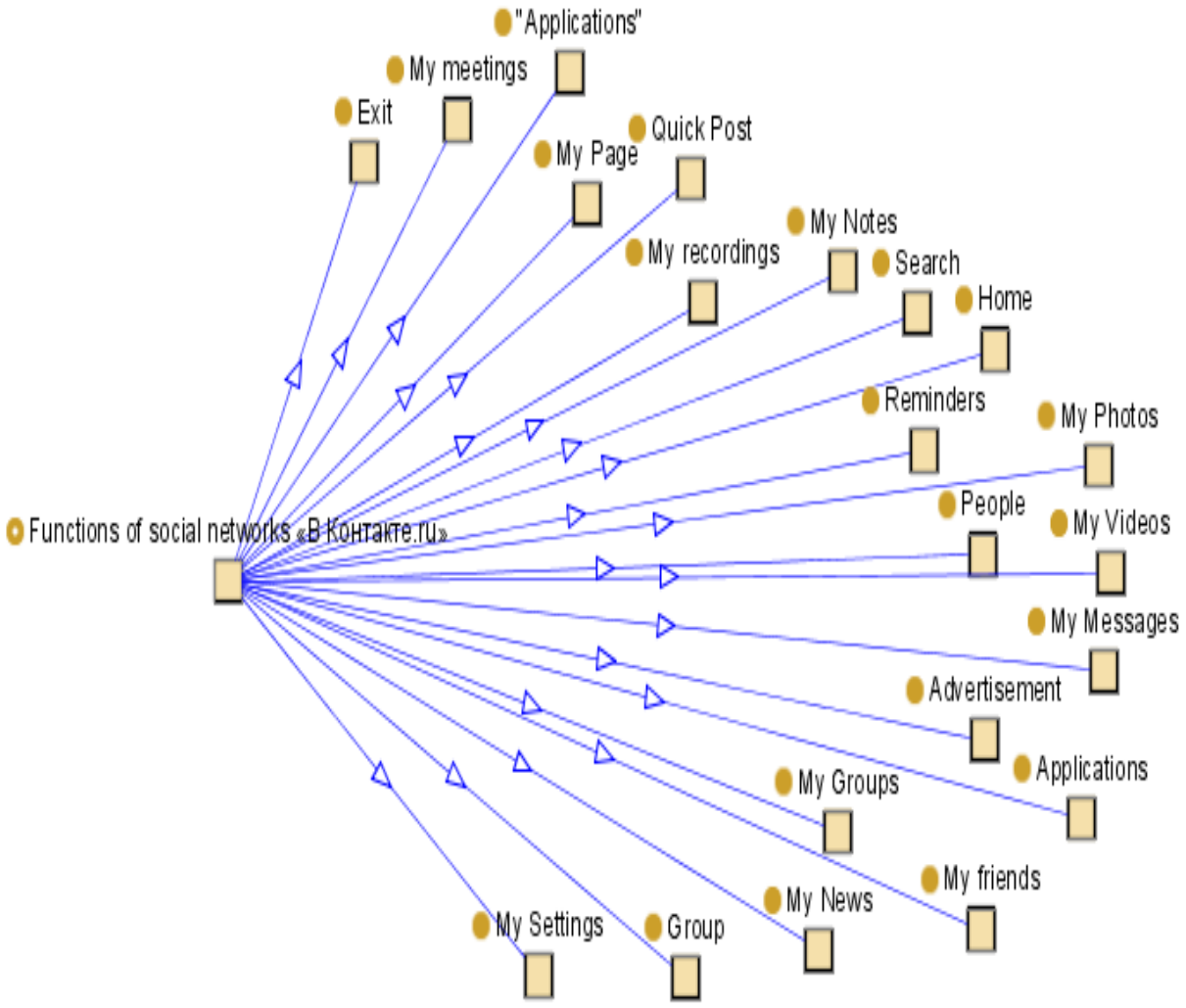


Figure 1. The social network «В Контакте.ru» functions of the first classification level.

The recommended informative placement of the functions of the first level of the hierarchy, taking into account the systemological classification analysis use and the natural classification criteria, is shown in Figure 2. The functions placement is understood as their placement in the networks workspace. The informative placement is understood as the functions structure their hierarchy in the social network menu. Our informative placement displays the functions relationship taking into account the knowledge systematization and the relations semantics between them in the best possible way.

Resulting from the analysis it was proposed to divide all the functions of the first level of the hierarchy on such groups: user information, my data, my messages, search and the function of input/output from the network workspace. Such functions placement will allow reducing the load on the user and speeding up the process of the needed function search, as it reflects the functioning of the whole support ability relation. The substantial functions placement in the social networks proposed, as the result of the conducted functions system analysis, will facilitate the functions search for the social network new users, will reduce the sense loading on the user when working with the network, by reducing the number of functions on the same level of the hierarchy.

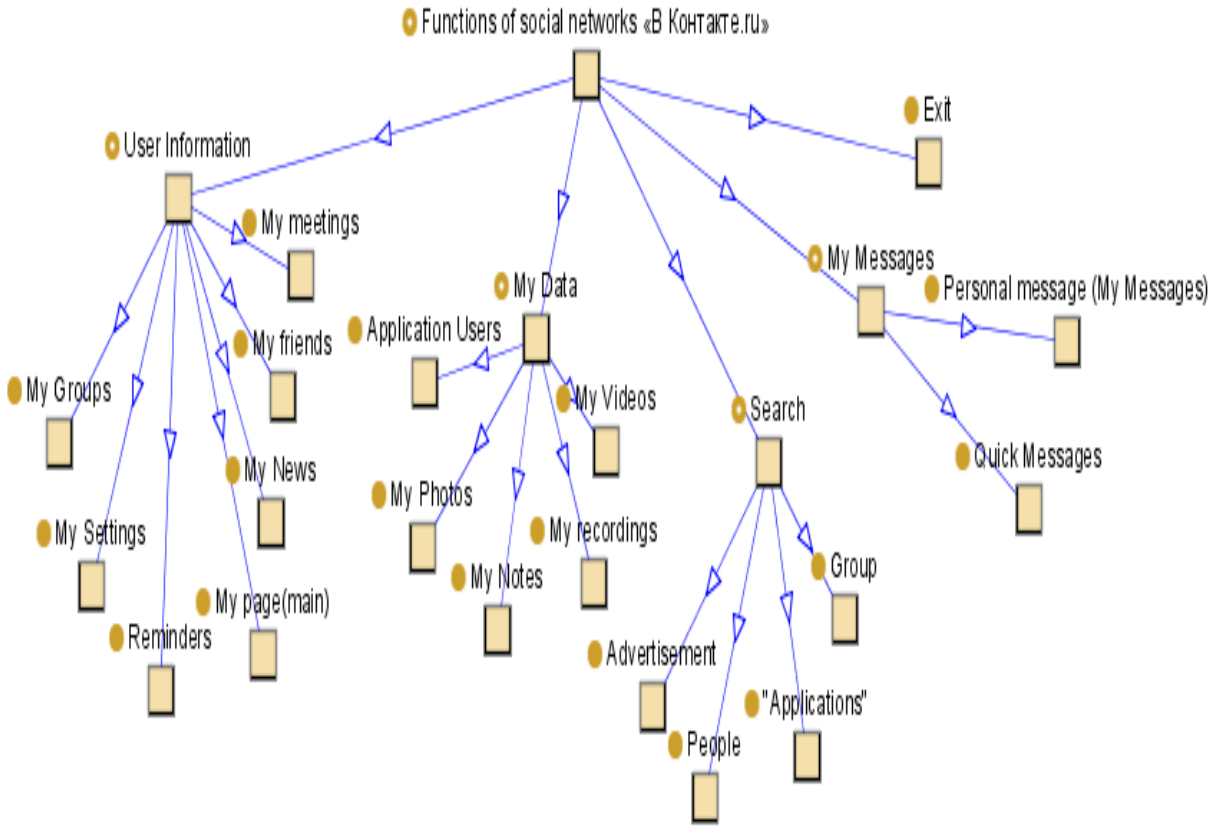


Figure 2. The social network "В Контакте.ru" systemological informative functions placement.

Also the knowledge systematization in the social networks specific functions will help to systematize the knowledge in the domain of the considered social networks in the Internet. After analyzing the social networks functions classification, created in the process of work, it was found that most social networks have a number of similar functions. In most cases, the functional destination coincided fully, and the functions differed significantly, for example, the function "Work with groups" appeared under the name "Groups", "Communities", "Groups and discussions".

The created functions classification reflects the semantics of the functions and of the relations between them, as well as the results of the knowledge systematization in the social networks in the Internet domain. Increasing the levels of the hierarchy in the functions classification, shown in Figure 2, allows finding the needed function faster, increasing the user convenience, accelerating the further functions menu development. Such analysis of the considered social networks functions allows creating the recommended functions classifications for each of them.

The advantage of the proposed classification of the social networks in Internet functions is that it includes the functions considered in popular social networks «В Контакте.ru» (http://vkontakte.ru), «Википедия» (http://ru.wikipedia.org/wiki), «Мой Мир» (http://my.mail.ru/mail), «Connect.ua», «МойКруг» (http://moikrug.ru), «Science-community.org».

For these networks the functions classifications *by the relation of the functional ability of the whole support* were created that has given the possibility to develop the recommendations or the meaningful placement of the menu functions of the social networks according to the requirements of systemology and formal logic. As an example, in Figure 3 our recommended classification of functions of the first level of hierarchy by the relation "part-whole" for the social network of scientists «Science-community.org», implemented in a software tool Protégé 3.2 is shown.

The systemological classification analysis application has allowed justifying the obtained classifications of social networks in Internet, to take into account the objects essential properties of them. This classification gives the possibility to detect and predict the objects properties by their position in the classification, i.e. from the viewpoint of the possibility to apply the classification not only as an effective practical tool but also as a tool of the theoretical analysis in the correspondent domain.

The use of the systemological classification analysis allows formulating recommendations for the hierarchical structure of functions implementation in the social network, for their meaningful placement in the menu in accordance with the created classification. Such natural placement will allow to reduce significantly the load on the user, will improve his work, networks and the principles of their functioning mastering.

The obtained classifications of the social networks in Internet functions allow to determine easy which class this or that concrete function of social networks refers to with which the user may meet while working with social networks in Internet. The greatest number of functions refers to the functions of "search" and "work with network resources," the functions of "communication" are also important. This classification of the social networks in

Internet functions can be viewed as a parametric (including the classification of properties) one, because the classes functionality are seen from their names. Resulting from the functions of various social networks research the functions classification fragment, shown in Figure 4 was built. The created classification fragment allows determining to which class refer the functions of the first level of the hierarchy of the social networks: «В Контакте.ru» (http://vkontakte.ru), «Википедия» (http://ru.wikipedia.org/wiki), «Мой Мир» (http://my.mail.ru/mail), «Connect.ua», «МойКруг» (http://moikrug.ru), «Science-community.org». The functions search was done by means of the practical use of a concrete function to verify its functionality. First the functionality for each concrete function was determined, and then the function appurtenance to the concrete class was determined. The obtained fragment of the classification "social networks functions" was realized in the software tool Protégé 3.2 is shown in Figure 4. This software tool was chosen due to a number of advantages [Shcherbak, 2008, etc.].

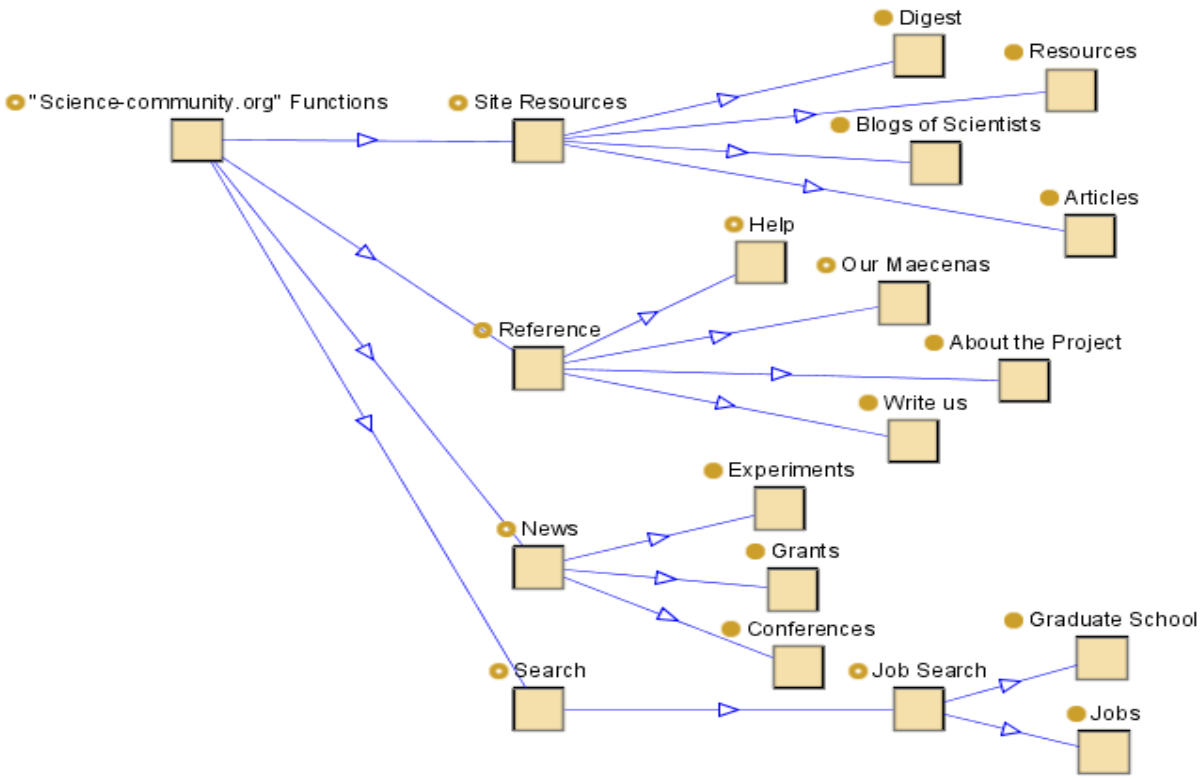


Figure 3. Recommended placement of the functions for the social network «Science-community.org»

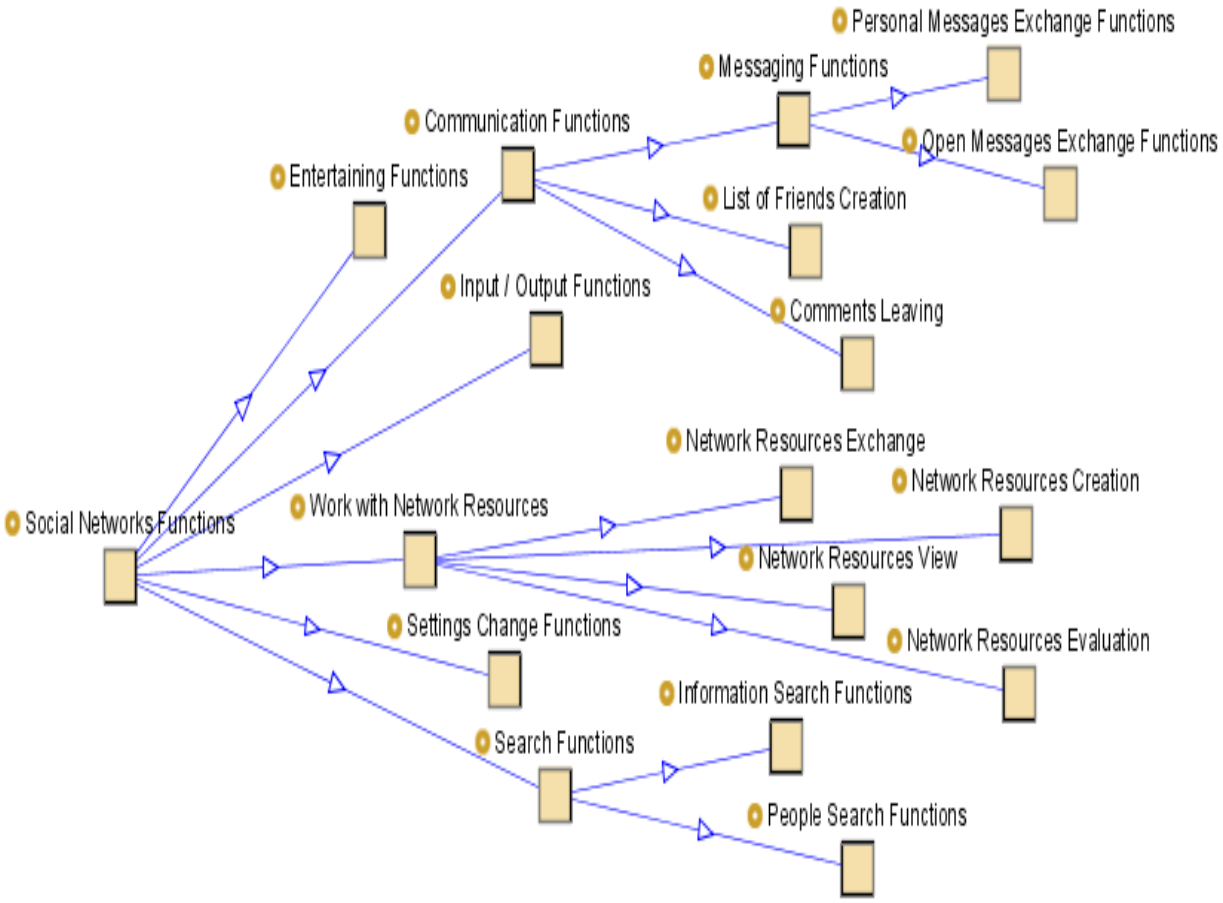


Figure 4. The social networks functions fragment classification by the relation of the functional ability of the whole support.

The obtained fragment of the social networks in Internet functions classification will allow becoming faster familiar with the functions of social networks in Internet, to choose more effectively the social network for registration, taking into account the functionality of the social network. The obtained results should be used for further knowledge systematization in the field of social networks in Internet.

The use of the method of systemological classification analysis allows obtaining new deep knowledge and systematizing knowledge in any domain in the most adequate and objective way taking into account the substantial properties and relations. The use of the systemological classification analysis allows evaluating the validity of any knowledge classification, the objects essential properties reflection in it; predicting new objects based on their properties.

Systemological research of social networks will allow systematizing knowledge in the social netwoks in the Internet domain and defining the appropriateness of various functions use in this or that social network, in a concrete organization.

Systemological Classification Analysis Application in the Social Networks Construction

Increase of the social networks in Internet influence of on the society has convinced many people to use social networks in business. Large corporations can afford to order a strong social network from firms of developers, but creation of such a network will require a lot of money. The enterprises (low-budget organizations) with a small income have not such a possibility, they may or attempt to use already functioning network or to attempt to create a social network by themselves. The latter variant is more advantageous, as the company itself regulates who will be the participant of the network, what tasks the social network must solve within the organization, etc. To create a social network in Internet it is necessary to use software for social networks creating.

Nowadays Internet is filled with a variety of software for the own social network creation. Many of them paid and (or) require deep knowledge in programming. There is also a number of software proposing to create a social network for free. This software proposes some free set of functions for a simple social network creating, there is also the possibility to use the supplement paid services.

The analysis of the software «Socialtext», «IBM Lotus Connections», «Jive SBS», «СвояСеть», «Connectbeam», «Ning», «Taba.ru» allows to make the conclusion that «Ning» (http://www.ning.com/), «Taba.ru» (http://taba.ru/), «СвояСеть» (http://svoyaset.ru/getform.html#) are the most acceptable for writing the recommendations to the social networks creation. They are conditionally free and do not require deep knowledge in programming. The disadvantage of the program service «Ning» is the absence of the interface in Russian. This disadvantage is significant for the recommendations to the social networks creation. In connection with it the software «Taba.ru», and «СвояСеть» were chosen. While creating the social network in «Taba.ru» it is recommended to use the social networks in Internet classification fragment shown in Figure 2.

In the process of writing recommendations the alternative menu creation of the social network has been tested using the systematological classification analysis. The social networks functions alternative menu created taking into account the results mentioned above was maximally approximated to the menu corresponding to the formal logic and systemological classification analysis. Unfortunately, the considered designers have the limited functionality and do not allow applying fully the results of the conducted research. In the process of work guidelines and recommendations to social networks creation in Internet in the software «Taba.ru», «СвояСеть» have been developed, the shortcomings and benefits of a social network creation in the selected designers have been revealed, as examples the demoversions of social networks in each of the designers have been created.

Model Creation of Knowledge Dissemination in an Organization with a Help of a Social Network

Despite the widespread of social networks in Internet, the models of knowledge dissemination in an organization by means of social networks sites have not been found. There is a number of articles describing the use of social networks for the social capital creating and using but there are no models of social networks implementation in an organization for knowledge dissemination and the employees' intellectual capital enhancing. The model creation of knowledge implementation and dissemination in an organization will allow increasing the organization competitiveness and solving such important practical tasks as the social networks implementation process acceleration, improving their functioning effectiveness and facilitating the process of knowledge acquisition and dissemination in the social network space by the employees.

When choosing a social network it is necessary to take into account several factors, such as the creation goals, the project budget, the tasks which will be solved by means of the social network, the expected users' range. The informational business model of a social network choice, aimed at the concrete organization problems solving, will help to facilitate the process of choosing a social network.

Systemology and the systemological classification analysis on the basis of Natural Classificaton application in social networks will allow increasing the functioning effectiveness of the fuctional menu and the functioning effectiveness of social networks, the networks implementation, facilitating the new functions implementation. Using the knowledge obtained during the social networks functions classification creation, the informational business model(Figure 5) describing the process of a social network in the Internet choice and development for increasing the organization competitiveness. This model realised in BPWin and describes the main processes in the organization when choosing and creating a social network (the definition of the goals and tasks solved by the social network; the means and the software tool for creating the social network choice; a brief description of the processes associated with the immediate introduction of a social network in operation, of the ways of promoting the social network use by the employees).

In the future it is planned to develop the business model of social networks application in the Internet for knowledge management. The model will include the methods of exchanging both explicit and tacit knowledge of knowledge that will allow to increase the effectiveness of the social networks sites application in the organization for knowledge management.

The use of the created informational business model will allow facilitating and accelerating the Internet social networks choice and implementation process in the company and minimizing changes necessary for the effective functioning of the social network in the Internet; will allow reducing costs during the social network in the Internet creation and use.

The results of the work can be used as recommendations for the construction or choice of a social network by the organization for increasing its competitiveness, strengthening the relationships between the employees

(increasing the social capital), increasing the intellectual capital of the employees and of the company on the whole.

The proposed results of the social networks may be used in the process of a learning organization creation, for decision making, intelligence technologies and artificial intelligence development.

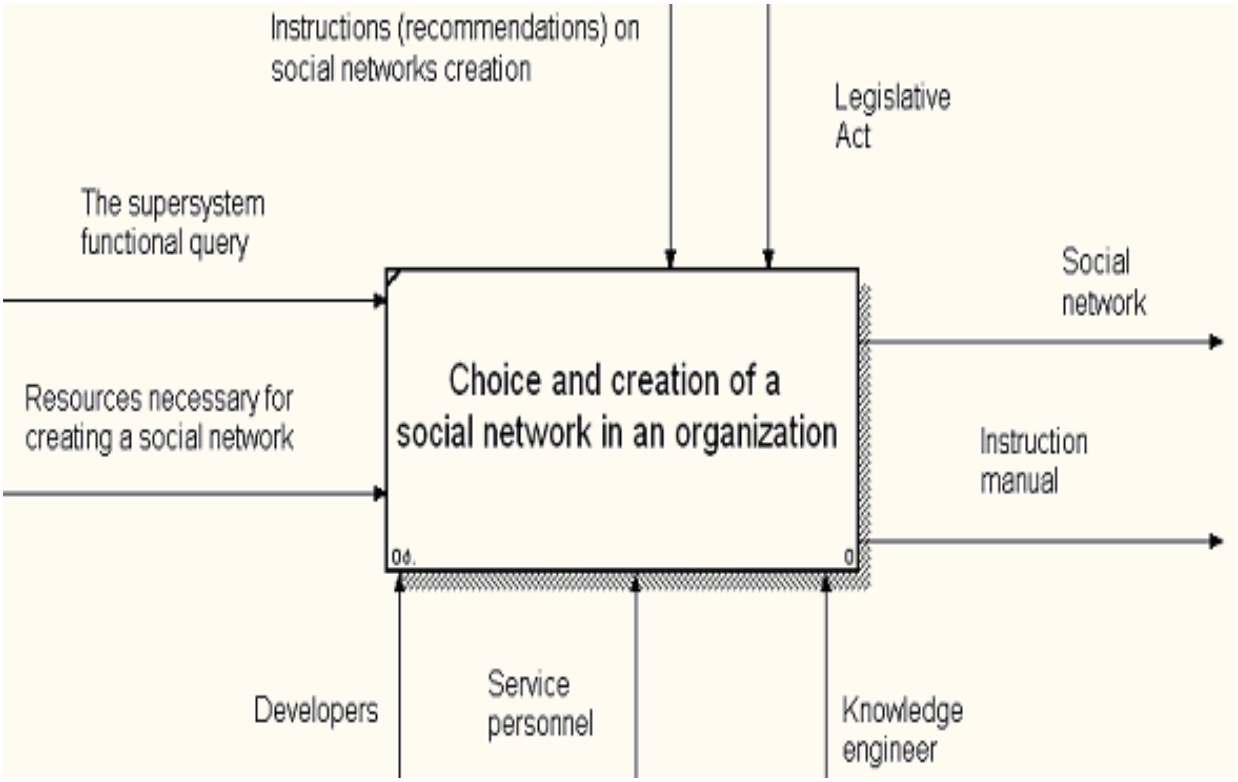


Figure 5. Context diagram of the informational model of the social network selection and creation in the organization.

Conclusion

The classifications of concepts are the basis of each science and are applied for solving various scientific-practical tasks. Now the classifications has got "the second birth" and are an main element of ontologies, computer models of knowledge, object-oriented analysis and modeling, intelligence technologies, knowledge management, decision making support and artificial intelligence, etc. That is why the role and the necessity of "good" classifications of concepts have increased now even more. Systemology application has allowed synthesizing system and classification analysis, discovering new criteria of systematics (natural classification) and their applying for knowledge systematization in any domain.

The Natural Classification criterion has been successfully used during the new method of the systemological classification analysis application. The results of the systemological research partially included in the paper may be used for the further knowledge systematization, creation of more effective alternative menus, natural language processing, etc.

Acknowledgements

The paper is partially financed by the project **ITHEA XXI** of the Institute of Information Theories and Applications FOI ITHEA and the Consortium FOI Bulgaria (www.ithea.org, www.foibg.com).

Bibliography

[[E. A. Solovyova, 1991] E. A. Solovyova. Mathematical Modeling of Conceptual System: a Method and Criteria of a Natural Classification (Article). New York: Allertion Press, Inc., V. 25, No. 2., 1991.

[E. A. Solovyova, 1999]. E. A. Solovyova. Natural Classification: Systemological Bases [In Russian], Kharkov: KhNURE, 1999.

[E. A.. Solovyova, 2000]. Mathematical and Systemological Foundations of Natural Classification (Article). New York: Allerton Press, Inc., V. 33, No. 4., 2000.

[E. A. Solovyova, et al, 2000] D.B. Elchaninov, S.I. Matorin]. Application Of Categories Theory To Research and To Modeling Of Natural Classification (Article). New York: Allerton Press, Inc., V. 33, No. 2., 2000.

[Bondarenko et al, 1998] M. F. Bondarenko, E. A. Solovyova, S. I. Matorin. Foundations of Systemolgy, [In Russian], Kharkov : KhTURE, 1998.

[M. Auxilio, M. Nieto, 2003]. M. Auxilio. An Overview of Ontologies: Technical Report, Mexico: Center for Research in Information and Automation Technologies, 2003.

[Byelyenkiy, 2008]. A. Byelenkiy Business Perspectives of Social Networks // http://www.compress.ru/article.aspx?id=18650&iid=865 [In Russian].

[Kuzmenko, 2009] Kuzmenko. Social Network // http://www.itpedia.ru/index.php/Социальная_сеть [In Russian].

[Shcherbak, 2008] S. S. Shcherbak. A Few Words about the Protocol Open Knowledge Base Connectivity (OCBC) and about the ontologies redactor Protégé // http://semanticfuture.net/index.php?title [In Russian].

[Danilov, 2010] A. Danilov To the question of the knowledge systematization in the social networks domain, Newsletter of National Technical University "KhPI", [In Russian], Kharkov: NTU "KhPI", 2010. - №67.-199 p.

Authors' Information

Mikhail Bondarenko *– Rector of Kharkov National University of Radio Electronics, Corresponding Member of the National Academy of Sciences of Ukraine, Lenin Ave., 14, Kharkov, 61166, Ukraine; e-mail: rector@kture.kharkov.ua*

Major Fields of Scientific Research: Intelligence Technologies, Information and Knowledge Management, System Analysis, Artificial Intelligence, Decision Making, Knowledge Research and Application, Natural Language Processing, Knowledge and Natural Language Modeling, Business Intelligence, Competitive Intelligence, Modern (e-) Learning, Knowledge-Based Systems and Technologies, Systemological Analysis.

Kateryna Solovyova *- Chief of Social Informatics Departmen and Knowledge Management Center, Professor, Doctor of Technical Sciences, Kharkov National University of Radio Electronics, Lenin Ave., 14, Kharkov, 61166, Ukraine; e-mail: si@kture.kharkov.ua*

Major Fields of Scientific Research: Knowledge Classification, Systematization, Elicitation, Acquisition and Modeling, Knowledge Management, Ontological Engineering, Systemological Analysis, Knowledge Research and Application, Decision Making, Knowledge-Based Systems and Technologies, Artificial Intelligence, Business Intelligence, Modern (e-) Learning, Competitive Intelligence, Cognitive Modeling, Intellectual Capital .

Andrey Danilov *– Social Informatics Department, Kharkov National University of Radio Electronics, Engineer, Lenin Ave., 14, Kharkov, 61166, Ukraine; e-mail: si@kture.kharkov.ua, Skil06@ukr.net*

Major Fields of Scientific Research: Social Networks, Ontological Engineering, Competitive Intelligence, Decision Making, Intelligence Technologies, Knowledge Research and Application, Knowledge Management, (e-) Learning, Artificial Intelligence, Systemological Analysis, Social Capital.

МЕТОДИКА ИСПОЛЬЗОВАНИЯ СРЕДСТВ СТРУКТУРИЗАЦИИ УЧЕБНОГО МАТЕРИАЛА

Сулима Е.Н., Миленин В.М

Аннотация. *Описываются средства структуризации информационных массивов на основе использования растущих пирамидальных сетей. Приводятся примеры обработки информации на сайте ОСТРОВ ЗНАНИЙ.*

Ключевые слова: *структуризация естественно-языкового текста, представление и обработка знаний.*

ACM Classification Keywords: *I.2 ARTIFICIAL INTELLIGENCE - I.2.4 Knowledge Representation Formalisms and Methods, K.3 COMPUTERS AND EDUCATION - K.3.1 Computer Uses in Education*

Введение

Создание формальной структуры учебного материала выражается в определении системы отношений между понятиями: меронимии, гипонимии, исключения, дополнения, ассоциации, сходства, различия, близости, отдаленности, причины, следствия, до и после и т.д. Связи устанавливаются не только между понятиями, но и между предметами, явлениями. Основной вывод исследований по запоминанию учебного материала тривиален и состоит в том, что структурированный материал заучивается легче, чем бессвязный. Чем больше связей нового материала со старым, тем легче осуществляется процесс запоминания. Запоминание информации не следует понимать как процесс "раскладывания" ее по заранее заготовленным "полкам". Одновременно с восприятием информации осуществляется как структурирование информации, так и структурирование самой памяти под влиянием воспринимаемой и уже хранящейся информации. Воспринятая информация отображается в структуре памяти.

Какие же типы компьютерных информационных моделей наиболее близки к моделям, которые использует человек для осуществления своей жизнедеятельности? В информационно-аналитических системах для обработки сложных структур разнотипных данных и знаний используются логико-лингвистические модели [Поспелов, 1981], т.е. такие модели, в которых основными элементами является не числа и вычислительные операции, а имена и логические связи. Естественно-языковый текст также может быть адекватно представлен с помощью логико-лингвистической модели.

Одной из хорошо апробированных возможных реализаций логико-лингвистических моделей является организация памяти информационно-аналитических систем в виде *растущих пирамидальных сетей* (РПС) [Гладун, 2000].

Построение РПС

Процесс построения растущей пирамидальной сети описан во многих публикациях [Гладун, 1994, Гладун, 1987]. Напомним основные понятия и процессы, происходящие при построении РПС. *Растущей пирамидальной сетью* называется ациклический ориентированный граф, в котором нет вершин, имеющих одну заходящую дугу. Примеры РПС приведены на рис.1,2,3. Вершины, не имеющие заходящих дуг, называются *рецепторами*, остальные вершины – *концепторами*. Подграф пирамидальной сети, включающий вершину *a* и все вершины, от которых имеются пути к вершине *a*, называется *пирамидой* вершины *a*. Вершины, входящие в пирамиду вершины *a*, образуют ее *субмножество*. Множество вершин, к которым имеются пути от вершины *a*, называется ее *супермножеством*.

В субмножестве и супермножестве вершины *a* выделяются 0-субмножество и 0-супермножество, состоящие из тех вершин, которые связаны с вершиной *a* непосредственно. При построении сети входной информацией служат наборы значений признаков, описывающих некоторые объекты произвольной предметной области (материалы, состояния агрегата, ситуации, термины учебного курса и т.п.). Рецепторы соответствуют значениям признаков. В различных задачах это могут быть имена свойств, отношений, состояний, действий, объектов или классов объектов. По мере ввода в сеть новых объектов формируются концепторы – специальные вершины, соответствующие пресечениям множеств рецепторов различных объектов, а также описаниям объектов в целом. В начальном состоянии сеть состоит только из рецепторов. Концепторы формируются в результате работы алгоритма построения сети. Приведем описание алгоритма построения сети[Гладун, 1994].

При вводе признакового описания объекта, рецепторы, соответствующие значениям признаков, входящим в описание, переводятся в *состояние возбуждения*. Процесс возбуждения распространяется по сети. Концептор переводится в состояние возбуждения, если возбуждены все вершины его 0–субмножества. Рецепторы и концепторы сохраняют состояние возбуждения в течение выполнения всех операций достройки сети.

Пусть при вводе описания объекта F_a – подмножество возбужденных вершин 0-субмножества вершины *a*; *G* – множество возбужденных вершин сети, не имеющих других возбужденных вершин в своих супермножествах.

Ввод новых вершин производится по следующим правилам.

Правило А1.

Если вершина *a* не возбуждена, и множество F_a содержит более одного элемента, то дуги, соединяющие вершины из множества F_a с вершиной *a*, ликвидируются, и в сеть вводится новый концептор, который соединяется заходящими дугами с вершинами множества F_a и исходящей дугой с вершиной *a*. Новая вершина находится в состоянии возбуждения.

Выполнение правила А1 иллюстрируется рис. 1 (а, б, в). Сеть рис. 1 (б) возникает после возбуждения в сети рис. 1 (а) рецепторов 4, 5, а сеть рис. 1 (в) возникает после дополнительного возбуждения в сети рис. 1 (б) рецепторов 2, 3.

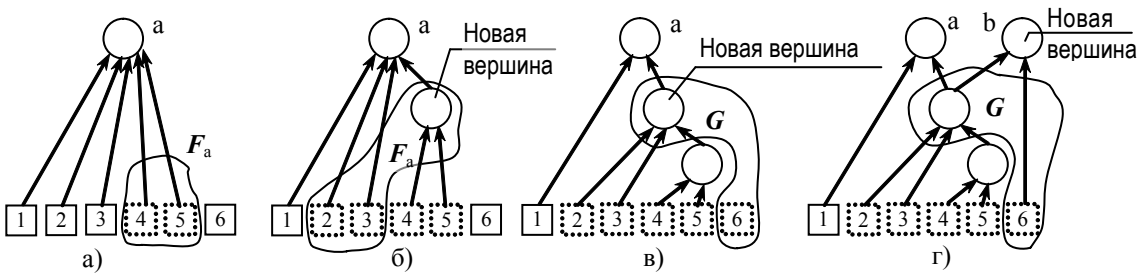


Рис. 1.

Как следует из правила А1, условием ввода в сеть новой вершины является ситуация, когда некоторая вершина сети оказывается не полностью возбужденной (возбуждаются не все, но не менее двух вершин ее 0-субмножества). Новые вершины вводятся в субмножества не полностью возбужденных вершин. Они представляют в сети пересечения описаний объектов.

После введения новых вершин во все участки, где удовлетворяется условие правила А1, выполняются правила А2 или А3, завершающие построение пирамиды объекта.

Правило А2.

Если множество *G* содержит более одного элемента и не включает вершины, помеченной именем введенного объекта, к сети присоединяется новый концептор, который соединяется заходящими дугами со всеми вершинами множества *G*. Новая вершина находится в возбужденном состоянии.

Выполнение правила А2 иллюстрируется рис.1 (в, г). Сеть рис.1 (г) возникает после возбуждения в сети рис.1 (в) рецепторов 2,3,4,5,6.

Особенности РПС

Далее остановимся на особенностях РПС, которые позволяют эффективно их использовать для решения задач. В РПС структурирование памяти осуществляется путем построения иерархической сетевой структуры одновременно с восприятием информации. Сформированная структура отображает особенности решаемой задачи. Алгоритм построения сети обеспечивает автоматическое установление ассоциативной близости между объектами на основе общих элементов из их описаний. Таким образом, при построении сети формируются конъюнктивные классы объектов, т.е. осуществляется классификация без учителя. При добавлении новых объектов нет необходимости выполнять операции поиска и преобразования сети по всем введенным объектам – формирование концепторов ограничено небольшими фрагментами сети и происходит только в пирамидах тех объектов, рецепторы которых совпадают с рецепторами нового объекта. Принципиальное решение проблемы сокращения времени

поиска в РПС состоит в отнесении операций по установлению сходства объектов на этап ввода информации об объектах в память, который выполняется только один раз при построении сети.

РПС удобны для выполнения операций ассоциативного поиска [Гладун, 1994, Гладун, 2000]. Например, можно выбрать все объекты, которые включают заданное объединение значений признаков, прослеживая пути, выходящие из вершины сети, которая отвечает этому объединению. Для выбора всех объектов, описания которых пересекаются с описанием заданного объекта, достаточно проследить пути, которые выходят из вершин пирамиды этого объекта. Все поисковые операции пирамидальной сети ограничиваются сравнительно малым участком сети, которая включает пирамиду объекта и вершины, непосредственно связанные с ней. Переход от конвергированных представлений объектов (концепторов) к развернутым - наборам рецепторов осуществляется путем просмотра пирамид объектов в различных направлениях (сверху вниз и снизу вверх). В результате появляется возможность решать практические задачи на основе анализа больших объемов данных.

Важным свойством пирамидальных сетей является их иерархичность, которая позволяет естественным образом отображать структуру составных объектов, родовидовые связи, связи объект-свойство. Обучение РПС состоит в формировании в них структур, представляющих понятие [Гладун, 1994]. Понятие – элемент системы знаний, представляющий собой обобщённую логическую признаковую модель класса объектов, с помощью которой реализуются процессы распознавания и генерации моделей конкретных объектов. После построения пирамидальной сети создается структура, являющаяся представлением описаний объектов. Каждая вершина сети определяет конъюнктивный класс объектов. РПС позволяет построить понятие в форме логического выражения, включающего имена рецепторов и логические операции дизъюнкции, конъюнкции и отрицания [Gladun, 2008]. Построенное понятие включает все наиболее важные признаки, характеризующие класс, и отображает характерные для этого класса логические связи между этими признаками. Представление понятия в виде логического выражения является наглядным, хорошо интерпретируемым и может быть использовано с целью более глубокого понимания закономерностей, присущих предметной области. Важной особенностью метода формирования понятий в РПС является возможность введения в понятие так называемых разделительных признаков, которые не принадлежат объектам исследуемого класса. В результате сформированные понятия имеют более компактную логическую структуру. РПС реализованы в программной системе "Конфор" [Гладун, 2002]. Система "Конфор" не является предметно-ориентированной и может быть использована для анализа данных в любой предметной области при условии, что данные представленные как множество описаний объектов в виде набора значений признаков.

Структуризация учебного материала с использованием РПС

Содержательная (семантическая) структура учебного материала предполагает смысловое, а не механическое запоминание. Такая структура отражает объективные зависимости, существующие в

природе, обществе и между людьми. При механическом запоминании материал заучивается таким, каким он предлагается, при смысловом - происходит организация и упорядочение сообщаемой информации, ее фильтрация. Современный школьник должен уметь в достаточной степени систематизировать и обобщать свои знания, знать соотношения и связи между структурными элементами системы знаний. Автоматизированная структуризация учебного материала позволяет в первую очередь в рамках электронного учебного ресурса связать контент ресурса в единую структуру, проследить связи между понятиями, используемыми в различных информационных источниках, которыми постоянно пополняется учебный ресурс.

Методика использования средств автоматизированной структуризации учебного материала включает следующие этапы:

1) автоматизированный поверхностный семантический анализ текста [Величко, 2009];

2) выделение терминов из текста на основе проведенного анализа [Величко, 2009];

3) выделение отношений между терминами на основе проведенного анализа;

4) для каждого проанализированного текста – формирование файла с выделенными терминами и отношениями между ними;

5) построение растущей пирамидальной сети, объединяющей термины и связывающей отдельные учебные материалы в единое информационное пространство.

Рассмотрим пример применения описанной технологии к обработке информации сайта «ОСТРОВ ЗНАНИЙ» (http://shkola.ostriv.in.ua). Проанализируем две статьи из раздела «Школа - Навчання - Додаткові матеріали»: «Будова Сонця» (Рис. 2), опубликована 03 Травня 2007 и «Класифікація зірок» (Рис. 3) опубликована 17 Січня 2007. Фрагменты списков терминов, автоматически построенных на основе лингвистического анализа статей приведены на Рис.4 и Рис.5. На Рис.6. приведен фрагмент построенной пирамидальной сети, объединяющей содержание двух учебных материалов. Используя сетевое представление терминологического содержания текстов легко проследить связи между терминами, связать в единое целое учебный материал, опубликованный в различное время. Графическое представление связей между понятиями улучшает качество восприятия учебного материала, позволяет определить место научного факта в обобщенной научной картине мира, обеспечивает формирование структурно-логических знаний учащихся. Навигация по сети позволяет переходить от одного факта к другому, находить места исходного текста, связанные с упоминанием того или иного факта. Автоматически сформированные термины позволяют уточнить списки ключевых слов, использующихся в описании статьи для последующего поиска по ключевым словам и рубрикации статей.

Будова Сонця

Зовнішній шар Сонця складається з фотосфери і хромосфери. Фотосфера, що є видимою поверхнею Сонця, має товщину приблизно 500 км і температуру близько 6000°К.

У фотосфері виявляється помітна активність в першу чергу у формі темних областей, так званих сонячних плям. За ними спостерігали ще в античні часи, але, не дивлячись на це, їх справжня природа не відома.

У пізніші часи Галілео Галілей "наново відкрив" сонячні плями, хоча це "відкриття" довго оспорювалося Крістофером Шейнером. Спостереження за сонячними плямами серед іншого, привело до краху арістотелівсько-птолемеєвської моделі Всесвіту, згідно якої всі зірки є ідеально неподільними сферами. Систематичні спостереження за сонячними плямами почалися приблизно в середині XVIII століття.

Сонячні плями здаються темними, але не тому, що вони дійсно чорні. Просто вони холодніші за фотосферу, що їоточує. Навколо найтемнішої зони плями, так званої тіні, є проміжна світла зона, так звана півтінь. Температура тіні рівна 4300-4800°К, тобто на 1000-1500°К нижча за температуру фотосфери. А ось температура півтіні рівна 5400-5500°К. Інтенсивність свічення в тіні складає приблизно 32% від інтенсивності свічення фотосфери, а півтіні - 80%. Схоже, що зниження температури усередині плям зв'язане сильними магнітними полями, відкритими Джорджем Еллері Хейлом в 1908 році. Такі поля заважають регулярному конвективному

Рис.2. Фрагмент учебного материала «Будова Сонця».

Класифікація зірок

При досягненні в надрах зірок високої густини і температури (біля 10-12 млн. К) починаються термоядерні реакції синтезу елементів - основне джерело енергії більшості зірок. Атмосфери зір - поверхневі шари зірок, у яких формується спектр спостерігаючого випромінювання зірок. В атмосфері зірок звичайно виділяють самий зовнішній шар: протяжну корону, потім хромосферу і розташовану ще глибше фотосферу. Маси зірок (M) знаходяться в межах від 0,04 до ~ 60 M☉, світності (L) від 0,5 до сотень тис. L☉.

Зірки класифікують по світності, масі, температурі поверхні, хімічному складу, особливостям спектру. На визначених етапах зіркової еволюції ряд зірок проходить через стадію нестаціонарності. У залежності від маси зірки наприкінці еволюції стають або **білими карликами**, або **нейтронними зірками**, або **чорними дірами**. Існують зоряні каталоги змінних зірок, подвійних зірок, пульсарів. Поділ зірок за їхньою світністю здійснюється на групи: **надгіганти, яскраві і слабкі гіганти, субгіганти, карлики, субкарлики і білі карлики**. Розходження світимості в надгігантів і білих карликів понад 20 зоряних величини (більш ніж у 108 разів). Діаграма Герцшпрунга-Рессела виражає зв'язок між світністю і температурою (спектральним класом чи показником кольору) зірок. На цій діаграмі близькі за фізичними властивостями зірки займають відособлені області: головну послідовність зірок, послідовності надгігантів, яскравих і слабких гігантів, субгігантів, субкарликів, білих карликів і ін. Головна послідовність діаграми Герцшпрунга-Рессела, вузька смуга на цій діаграмі, у межах якої знаходиться переважна більшість зірок, перетинає діаграму по діагоналі (від високих до низьких світимостей і температур). Зірки головної послідовності (до них, зокрема, відноситься Сонце) мають однакове джерело енергії термоядерної реакції кисневого циклу. Зірки знаходяться на головній послідовності протягом приблизно 90% усього часу зоряної еволюції. Цим пояснюється переважна концентрація зірок в області головної послідовності.

Рис.3. Фрагмент учебного материала «Класифікація зірок».

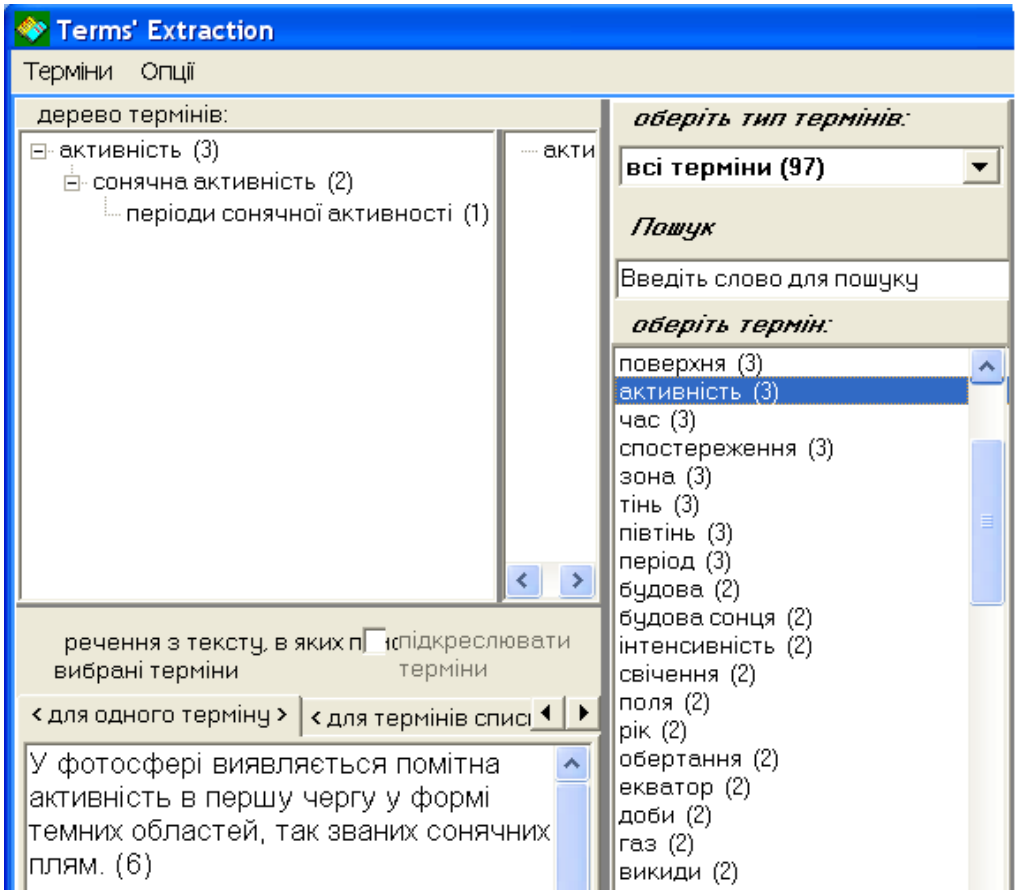


Рис.4. Фрагмент списка терминов, выделенных при анализе учебного материала «Будова Сонця».

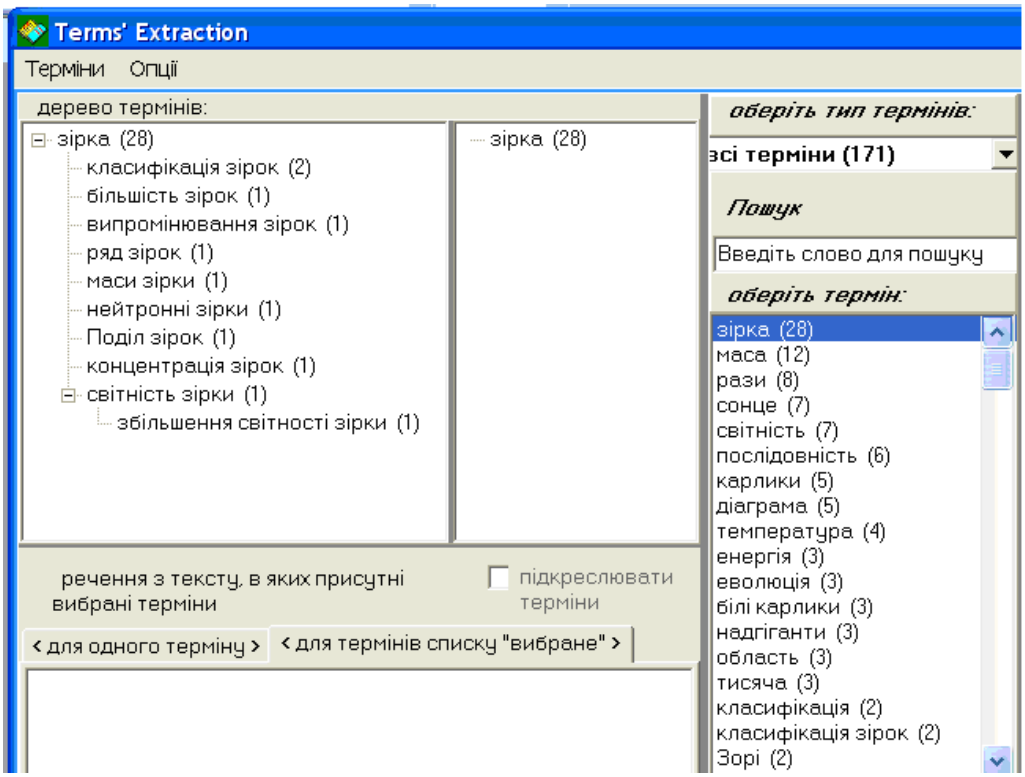


Рис.5. Фрагмент списка терминов, выделенных при анализе учебного материала «Класифікація зірок».

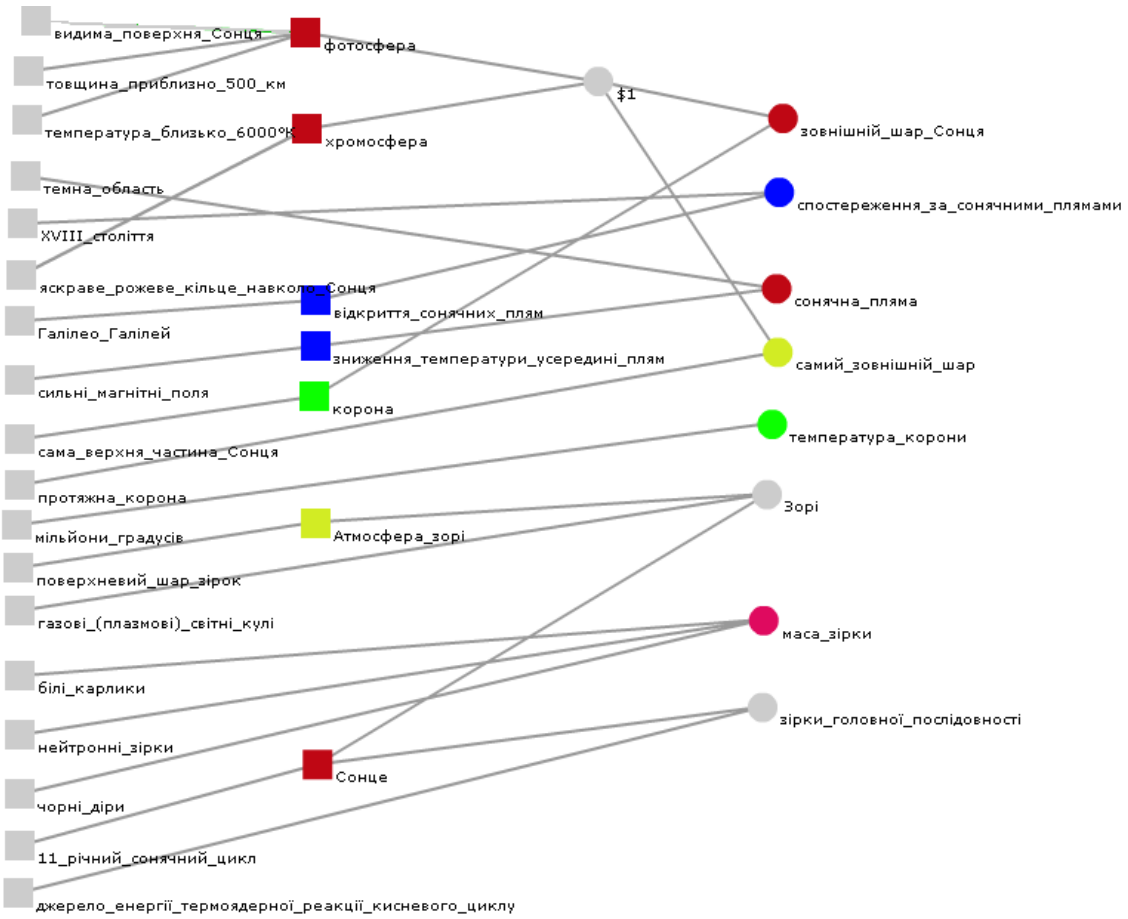


Рис.6. Фрагмент построенной пирамидальной сети, представляющей содержание двух учебных материалов

ЗАКЛЮЧЕНИЕ

Растущая пирамидальная сеть является сетевой памятью, самонастраивающейся на структуру входной информации. Такие качества, как простота внесения изменений, совмещение процессов ввода информации с ее классификацией, обобщением и выделением существенных признаков, высокая ассоциативность, делают растущие пирамидальные сети важным компонентом интеллектуальных систем. Одной из главных задач логической структуризации учебного материала, как средства систематизации и обобщения знаний учащихся, является формирование не формальных, а осознанных знаний, позволяющих их активное применение в творчестве, решении задач анализа, синтеза, обобщения, поиска новых знаний. Современные средства автоматизированной структуризации учебного материала помогают решать задачу формирования активных знаний у учащихся, используя возможности системно-структурного метода для управления познавательной деятельностью и выделяя ориентиры для самостоятельного обновления и пополнения знаний.

Литература

[Поспелов, 1981] Поспелов Д.А. Логико-лингвистические модели в системах управления.-Москва: Энергоиздат.- 1981.

[Гладун, 2000] Гладун В.П. Партнерство с компьютером. Человеко-машинные целеустремленные системы.-Киев: Port-Royal. – 2000.

[Гладун, 1994] Гладун В. П. Процессы формирования новых знаний. – София: СД «Педагог 6». – 1994.

[Гладун, 1987] Гладун В.П. Планирование решений.-Киев: Наукова думка.– 1987.

[Gladun, 2008] Victor Gladun, Vitaly Velichko, Yurii Ivaskiv. Selfstructurized Systems. International Journal "Information Theories & Applications". FOI ITHEA, Sofia. – Volume 15 – 2008, Number 1 – P. 5-13.

[Гладун, 2002] Гладун В.П., Ващенко Н.Д., Величко В.Ю. Программный комплекс решения аналитических задач методами правдоподобного вывода // Искусственный интеллект.- 2002.-№2.- с.430-437.

[Величко, 2009] Виталий Величко, Павел Волошин, Светлана Свитла Автоматизированное создание тезауруса терминов предметной области для локальных поисковых систем. "Knowledge – Dialogue – Solution" International Book Series "INFORMATION SCIENCE & COMPUTING", Number 15. – FOI ITHEA Sofia, Bulgaria. - 2009. – pp.24-31.

About:

ITHEA® International Scientific Society

ITHEA International Scientific Society (**ITHEA ISS**) is aimed to support growing collaboration between scientists from all over the world.

ITHEA ISS a successor of the international scientific co-operation organized within 1986-1992 by international workgroups (**IWG**) researching the problems of databases and artificial intelligence. As a result of tight relation between these problems in 1990 in Budapest appeared the international scientific group of Data Base Intellectualization (**IWGDBI**) integrating the possibilities of databases with the creative process support tools. The leaders of the IWGDBI were Prof. Victor Gladun (Ukraine) and Prof. Rumyana Kirkova (Bulgaria).

Starting from 1992 until now the international scientific co-operation has been organized by the Association of Developers and Users of Intellectualized Systems (**ADUIS**), Ukraine. It has played a significant role for uniting the scientific community working in the area of the artificial intelligence.

To extend the possibilities for international scientific collaboration in all directions of informatics by wide range of concrete activities, in 2002 year, the Institute for Information Theories and Applications FOI ITHEA (**IITA FOI ITHEA**) has been established as an international nongovernmental organization. IITA FOI ITHEA is aimed to support international scientific research through international scientific projects, workshops, conferences, journals, book series, etc. The achieved results are remarkable. IITA FOI ITHEA became worldwide known scientific organization. One of the main activities of the IITA FOI ITHEA is building the ITHEA International Scientific Society aimed to unite researches from all over the world who are working in the area of informatics.

Till now, the **ITHEA International Scientific Society** was joined by more than **2600** members from **48** countries all over the world: *Armenia, Azerbaijan, Belarus, Brazil, Belgium, Bulgaria, Canada, China, Czech Republic,*

Denmark, Egypt, Estonia, Finland, France, Germany, Greece, Hungary, India, Iran, Ireland, Israel, Italy, Japan, Jordan, Kyrgyz Republic, Latvia, Lithuania, Malaysia, Malta, Mexico, Moldova, Netherlands, Peru, Poland, Portugal, Romania, Russia, Scotland, Senegal, Serbia, Montenegro, Sweden, Spain, Sultanate of Oman, Turkey, UK, Ukraine, and USA.

ITHEA Publishing House (**ITHEA PH**) is the official publisher of the works of the ITHEA ISS. The scope of the books of the ITHEA ISS covers the area of Informatics and Computer Science. ITHEA PH welcomes scientific papers and books connected with any information theory or its application.

ITHEA ISS has two International Journals, established as independent scientific printed and electronic media, published by ITHEA Publishing House:

- **International Journal "Information Theories and Applications"** (**IJ ITA**), since 1993
- **International Journal "Information Technologies and Knowledge"** (IJ ITK) since 2007

Every year, ITHEA ISS organizes many scientific events called **"ITHEA ISS Joint International Events of Informatics" (ITA**). For the first 10 years of the XXI century, ITHEA ISS has organized **64** International Conferences and workshops and has published **44** scientific books.

All journals and books of ITHEA ISS are edited by the Institute of Information Theories and Applications FOI ITHEA, in collaboration with the leading researchers from the: *Institute of Cybernetics "V.M.Glushkov", NASU* (Ukraine)*; Institute of Mathematics and Informatics, BAS* (Bulgaria)*; Institute for Informatics and Automation Problems, NAS of the Republic of Armenia; University of Hasselt* (Belgium)*; Natural Computing Group (NCG) of the Technical University of Madrid* (Spain)*; Astrakhan State Technical University* (Russia)*; Taras Shevchenko National University of Kiev* (Ukraine)*; University of Calgary* (Canada)*; VLSI Systems Centre, Ben-Gurion University* (Israel).

The great success of ITHEA International Journals, Books, and Conferences belongs to the whole of the ITHEA International Scientific Society.

More information may be obtained from www.ithea.org.

CONTENT OF IJ ITA, VOL. 17, 2010

Table of contents of IJ ITA Vol. 17, No.: 1

Table of contents of IJ ITA Vol. 17, No.: 2

Table of contents of IJ ITA Vol. 17, No.: 3

Table of contents of IJ ITA Vol. 17, No.: 4