ISSN 0005-1055, Automatic Documentation and Mathematical Linguistics, 2019, Vol. 53, No. 3, pp. 143–159. © Allerton Press, Inc., 2019.
Russian Text © The Author(s), 2019, published in Nauchno-Tekhnicheskaya Informatsiya, Seriya 2: Informatsionnye Protsessy i Sistemy, 2019, No. 6, pp. 28–44.

AUTOMATION
OF TEXT PROCESSING

Method of Coding for Multicomponent Objects (RHA) and Its Application for Ordering Roman Fonts

E. T. Petrova[a, *], T. G. Petrov[b, **], S. V. Chebanov[c, *], and S. V. Moshkin[d, ****]**

[a]*Graphic artist, independent researcher, St. Petersburg, Russia*
[b]*Sokolov Co Ltd, St. Petersburg, Russia*
[c]*St. Petersburg State University, St. Petersburg, 199034 Russia*
[d]*National Library of Russia, St. Petersburg, 191023 Russia*
**e-mail: katia.petrova@gmail.com*
***e-mail: tomas_petrov@rambler.ru*
****e-mail: s.chebanov@gmail.com*
*****e-mail: svmoshkin52@gmail.com*

Received December 27, 2018

Abstract—The authors propose to systematize the font frameworks of the "H" sign types of different fonts on the base of the *RHA* information language, which has been developed for different types of objects, by assigning a code to the framework by placing the H font in the standardized window (FontWindow), with the subsequent calculation of quantity characteristics (characterizations). The codes are rendered in a diagram. The sorting of codes is carried out using the SBCO special alphabet, where S, B, C, O are, on one hand, code elements, and on the other hand, component designations of fonts frameworks. The sorting of codes is carried out by alphabets (1) ratings of assignment of components fields in a FontWindow, (2) entropies, and (3) anentropies. The codes are accompanied by the names of coded fonts. The authors propose a principle of placing codes, which does not depend on the name, style, width, the purpose of the font, and its author. The described method organizes H fonts by resembling of their frameworks. The created list-catalog can include all Latin and Cyrillic H fonts of direct inscriptions. The catalog layout includes 99 codes of H font frameworks ranged into 6 classes from 24 possible classes in the system made by authors. The proposed variant of the applying of the method can be used as a base of creating of a common method of coding and systematization of bicolored (in the present paper) and multicolored images on the plane, including maps and other presentations, which include sets of components of different colors and forms.

Keywords: font sign, allograph, FontWindow, letter frame, alphabetical ordering of font codes, letter frame entropy, letter frame anentropy, *RHA* code of the frame of sign type, *RHA*-system of the frames of sign types, *R*-class of the frames of sign type, entropy–anentropic diagram, fonts-synonyms, font-clone, geographical maps, geological maps, multi-color images

DOI: 10.3103/S0005105519030087

INTRODUCTION

Font diversity, which arises as a result of a special type of artistic creativity, is represented by numerous variants, and each year the font and design studios release new ones. There are over 13 000 fonts. Designers, authors, and printers are often faced with the need to search for fonts of a desired pattern, while sorting the font database on a computer or on the website of the seller or font manufacturer. Pirated clone fonts exist that completely copy the design of the author's versions under a modified name [1, 2], synonym fonts created in the image of a basic font and similar to the base font, as well as fonts with the same name but a different pattern.

There are several ways to classify fonts [3], from which groupings can be distinguished by: historical aspects [4, 5], design features (with serif and sans serif [6, 7]), type of creation (handwritten, stencil, etc. [8, 9]), purpose (type, accidental, and character [10–12]). All of these methods for organizing fonts require expert judgment. Universal (today) is a simple ordering using a natural font alphabet, as used in programs (for example, *FontExplorer* and *FontExpert*), and in printed catalogs [13, 14], and on font sites [15–17]. However, the names (titles) of fonts do not have a semantic load; therefore when using them, without knowing the font "in person," without having measurable characteristics of the fonts and their ordering system, similar fonts can be found only with special applications [18]. One popular application [19] recognizes isolated characters of the Latin font supplied to the program input and offers several types of font names to choose from. The site *Identifont*, which offers to search for the desired font by answering questions, narrows the search to several font names [20]. These resources are available and easy to use; however, they are fragmented and do not give an unambiguous ordering of fonts according to their form among those similar in shape, they have no way of tracking the variability of the form from font to font. These applications do not include unlicensed fonts [18]. From a practical point of view, the current situation is quite acceptable for designers engaged in fonts; however it is impossible to see the general principles of organization of font variety, the tendency of their appearance, the specifics of their use, the features of the psychophysiology of perception, or to predict the appearance of new fonts with certain characteristics.

In [21] it was noted that the construction of an all-encompassing detailed classification of fonts is hardly possible at all, since they are the product of artistic creation. At the same time, when describing them, one usually seeks to cover the object in detail. However, if one does not require taking ***all*** the features of the objects that are being described into account, perhaps smoothly changing and not having definite boundaries between clearly different ones, these objects can be described, for example, by selecting some measurable properties that are important for a given group of objects and comparing objects via these properties, as natural scientists do (more precisely, we are talking about different ways of working: with idiography, that is, revealing diversity, and nomothetic, revealing laws; an example of their combination is [22]).

As a universal characteristic of continuously changing compositions of objects, a dual list of the names of the components that make up the object and the proportions of these components in the object that is subjected to entropy–anentropic analysis (the ***RHA*** method by T.G. Petrov [23]) was chosen. The analysis consists in that fact that after obtaining the rank formula R, C. Shannon's information entropy is computed ($H = -\Sigma p_i \ln p_i$, where p_i is the frequency of the i^{th} component normalized to 1 [24, p. 48]), which characterizes the uniform distribution of components, and anentropy ($A = -[(\Sigma \ln p_i)/n] - \ln(n)$, where n is the number of components that represent the object [24, p. 61]), as introduced by T.G. Petrov to characterize the non-uniformity of the contribution of the components to their distribution, based on the comparison of which meaningful conclusions are drawn in the studied objects. At the same time, H and A can be calculated for complete (which makes it possible to more fully represent their individuality) or truncated compositions of objects (for the purpose of their comparison, regardless of the nature of the objects and the ways of studying them).

The *RHA* [23] information language-method was developed for uniform description of this type of composition representation, which was originally proposed to describe the chemical compositions of rocks, and later manifested itself as a method of encoding and alphabetical ordering of the compositions of objects of any nature, the types of components of which are either discrete, or discretized [25, 26]. A modern discussion of the *RHA* method was given in [24–31]. The presentation of compositions (of minerals, rocks, texts, population ages, species composition of plankton, types of correspondence, etc.) processed by this method has two basic requirements: (1) certainty of the composition components to be specified, i.e., the uniqueness of their distinction and (2) the closeness of the sum of the sizes of the specified parts to 100% (or ***to unity***). The exact equality of the sum of the contents of the components to 100% is more likely to arouse suspicion in the quality of analysis than to indicate its accuracy. For the clause given in parentheses, it should be noted that in works on computer science, to which the *RHA* method belongs, percentages are not used as a measure of the intensity of a property or the frequency of events, but rather parts p_i of the whole taken as a unity ($\sum p_{ii} = 1$).

Thus, *RHA* is:

- a method of presenting data on composition;
- a method of primary processing of this data;
- an information retrieval language for the description of compositions;
- a method of coding data on the composition of a particular object;
- a code of the composition of this object;
- a method of ordering compositions.

Bearing in mind the diversity of the capabilities of the *RHA* method, it became natural to continue the development of the method using the example of encoding typographic fonts, which was started in [32].

The objectives of this article are to describe an alphanumeric method of encoding characters of fonts of a direct lettering, which allows encoding and ordering them by rank-and-entropy characterizations[1] in the form of a table, to present the catalog layout of the "H" sign types and to visualize their distribution on entropy–anthropic diagrams.

THE "H" SIGN TYPE AND THE STANDARD OF ITS DESCRIPTION

The frame of the "H" sign type (Fig. 1a) is understood as a schematic representation of the "H" sign, consisting of two vertical segments of comparable length, that is, boles and a connecting horizontal or slightly right-inclined connecting segment (with the left-inclined connecting segment, the Cyrillic "И" sign types begin, and with the right-inclined, "N"), whose height is not taken into account (except for the

[1] In contrast to the non-quantitative characteristics of objects, hereinafter, *quantitative characteristic*s, in our case, entropy and anentropy, will be called *characterizations*.

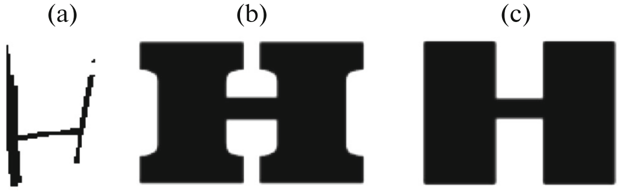


Fig. 1. Transitions from the sign type (a) to the H allograph of the *Braggadocio* (b) font and the "H" frame of the *Braggadocio* font (c).

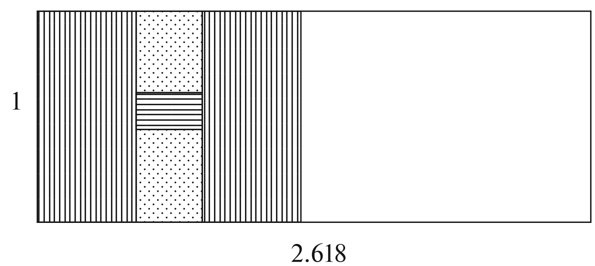


Fig. 2. The FontWindow with the "H" frame inscribed in it, where: ▥ S—boles; ⊟ B—connecting bar; ⊡ C—internal space of the allograph; □ O—residual space of the FontWindow.

prohibition of the occupation of the extreme upper and lower positions). This sign type, as common for the letter H of the Latin alphabet and Н of the Cyrillic alphabet, was chosen in [32] on this topic. This sign type, being a work of art in a particular font, may have numerous features that are not taken into account in a formalized description of the font. Thus, various modifications of the crossbar and boles of this type are ignored, as well as the unevenness of the tone or thickness of the components of the sign type, and we obtain a simplified image of the "H" sign type, that is, ***the frame of the "H" sign type***, which is subject to coding (Fig. 1). As a result, the "H" code of a specific font of a particular typeface will be obtained.

For comparability of different frames, their height is taken as a unity, and for compatibility of the width of frames, the maximum possible width is entered, for which the sum of the unity and the number φ of the golden proportion is taken, that is, the value 1 + 1.618 = 2.618. The addition of the unity is due to the fact that there are variants of "H", whose width exceeds 1.618 with a height equal to unity. For the visual presentation and comparability of the "H" frames of different fonts, a special window was built, that is, the FontWindow (F), with a fixed height, taken as a unity, and a width equal to 2.618. The resulting "H" frame of the *Braggadoci* font is placed in the FontWindow as an example (Fig. 2).

A FontWindow with an area of F_S with the frame placed in it has components for which we set the following notation:

(1) two main vertical bars of the frame, having a total area of S_S (*Stem*);

(2) a connecting horizontal bar of the frame with an area of B_S (Bar);

(3) an internal space of the frame with an area of C_S (*Counter*);

(4) a residual free space with an area of O_S (*Outdoor*).

The sum of the areas of the specified components is $F_S = S_S + B_S + C_S + O_S = 2.618$, which is taken as 100%, and in further calculations as 1.

We consider the idealized cases of the shape of the font character as the specified parameters approach the 100% of the area S_F of the FontWindow F (Fig. 3).

Therefore, we obtain a set of areas of the components of the frame of the letter and the area of free space in the FontWindow. This set will be converted to code by the *RHA* method.

REPRESENTATION OF THE FONTWINDOW WITH THE "H" FRAME IN THE *RHA* LANGUAGE

The Rank formula R

The description of sign types in the *RHA* language includes a semi-quantitative characteristic of the sign type by the composition of its components using the rank formula *R* [23] (the first symbol in the name of the method) and a quantitative description of the area composition of the sign type. The ranking of significances in describing sets of parts is a very common

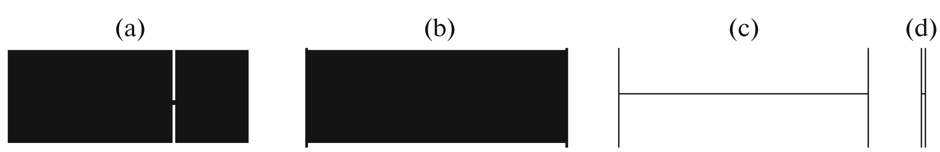


Fig. 3. Extreme–idealized forms of the H font sign:

(a) the form of H with the total area of boles S_S close to 100%. The greater the value of S_S is, the smaller the area of the internal space is and the wider the allograph is, the richer the font is;

(b) if the B_S area (connecting bar area) approaches 100%, a crossbar remains in the FontWindow with the disappearing participation of all other components of the allograph;

(c) approaching 100% of the internal space of the allograph (C_S) results in an approach to the zero areas of the remaining components in the FontWindow, which affects the lightness of the letter;

(d) with free area O_S approaching 100%, an allograph occurs with the disappearance of all other parts. The allograph is supernarrow.

procedure; however, such sequences have not been considered in science thus far, since they have not been understood as a particular object. However, rank formulas are a fundamentally new object of interest to informatics. The rank formula *R* of the component composition of the "H" sign type as the first part of the code is a sequence of names of the components of the frame for reducing the shares of their areas in the FontWindow. Thus, the rank formula contains two types of data about the frame: (1) qualitative characteristic about it, since it reports on a set of specific components of the frame located in the FontWindow and (2) ordinal characterizations, as the relative values of the components are reported in the rank formula: importance, significance, and the size of the frame parts in the FontWindow.

We construct the rank formula for the "H" frame of the *Braggadocio* font (see. Fig. 1 and *Appendix*[2], No. 32). The "H" frame of this font has the following areas of parts (in %): $S_S = 42.77$, $B_S = 0.30$, $C_S = 3.20$, $O_S = 53.73$. By writing the areas with their abbreviations on reducing the values and using inequality signs, we obtain the sequence O > S > C > B. After this, by eliminating inequality signs and values of areas, we obtain the rank formula $O_SS_SC_SB_S$. Another example is the frame of the *Distill* font (see *Appendix*, No. 66): $S_S = 7.84$, $B_S = 4.39$, $C_S = 36.50$, $O_S = 51.27$; it corresponds to the rank formula $O_SC_SS_SB_S$.

Thus, the proposed area ranking divides the possible variety of straight fonts into groups that have the same rank formulas. This is the first step in ordering the variety of fonts. An indefinitely large number of specific combinations of areas may correspond to the same rank formula. Among them are formulas that have neighboring components that practically do not differ in content. Equal signs are used to highlight such components.

There are no definite boundaries between the degrees of similarity from the "very" similar (two books of one edition) to "almost nothing in common" (two legs of a person and the cry of a gull). On the other hand, in some areas of knowledge (crystal chemistry and mineralogy) it turns out to be justified to consider values that differ by no more than 15 rel %[3] as equal.

In order not to look for other rationales and not stray from this topic, let us take this boundary in our case. Accordingly, we will put an equal sign between neighboring symbols in the rank formula of the frame areas if the result of dividing the previous (in the rank formula) area share by the next one does not exceed 1.15. In this case, the order of the components, that is, the frame parts, is preserved in the rank formula, and the order of these characters is given by the alphabet (S B C O) only if they are exactly equal.

The second part of the composition code, entered after the rank formula, the integral characterization of the frame, entropy (***H***), is ***mainly*** responsible for medium–large areas, the third, anentropy (***A***), is ***mainly*** responsible for small ones.

Entropy and anentropy allow one to more fully take the diversity of the set of sizes of areas into account. Let us turn to these concepts.

Entropy H

Let us consider the FontWindow with the H frame inscribed in it from the point of view of the degree of *equi*-dimensionality of the areas of its components: S_S, B_S, C_S, O_S. An extremely equi-dimensional complex is represented by a frame inscribed in the FontWindow, in which the areas of all components of the window contents are equal to each other.

An extremely *different*-sized content of the window is filled with a single, that is, any, part (Fig. 3). Accordingly, the number of extremely *different*-sized complexes will be equal to the number of components of the complex (i.e., four); the extremely equi-dimensional complex will be the only one that can be useful.

As a measure of uniformity, we take C. Shannon's information entropy, which is used in many branches of knowledge [33]. At the inception stage of the *RHA* method, it was proposed for geochemistry as a measure of the complexity of the chemical composition[4] [34]. It is determined by the formula:

$$H = -\sum p_i \ln p_i,$$

where p_i is the frequency of an event, in this case, it is equal to the shares of the areas of the components of the content of the FontWindow, whereby $\sum p_i = 1$. In the considered case, p_i is the share of the area occupied by the *i*th part of the content of the FontWindow. The value "$-p_i \ln p_i$" is the contribution of a single part of the FontWindow content to the entropy (Fig. 4). If $p_i = 0$, then the contribution of this component to the entropy is not calculated. At $p_i = 1$, the component contribution is 0.

As we see in Fig. 4, the dependences of the active[5] contribution to entropy and anentropy on the contents

[2] Hereinafter, in order to demonstrate examples of "H" sign types of different fonts, references to *Appendix* will be given before the principles of its construction are described.

[3] Relative percentages are understood as the ratio of two quantities, expressed as a percentage when dividing the greater one by the lesser one.

[4] The use of informational entropy as a measure of complexity is, in principle, not the only option for estimating complexity (for example, solving a complex task is not evaluated by entropy).

[5] Investments are qualified as active that directly depend on the presence of the considered components and do not contain explicit parameters related to the calculation procedure (according to the number of components). They are opposed to passive contributions, for which the calculation of the value explicitly depends on the parameters of the calculation. In our case, the parameter of the passive contribution in the formulas is equal to the number of components, four.

until the maximum contribution to the entropy is reached at $p = 0.368...$ behave symbatically, after reaching the specified maximum, in the opposite manner.

At uniform distribution of p, i.e., at $p_1 = p_2 = p_3 = ... = p_n = 1/n$ the entropy of the composition of the object, which can be interpreted as its complexity, is maximum and equal to $\ln n$ (natural logarithms), where n is the number of counted FontWindow components. Therefore, a filled FontWindow will be called *equi*-dimensional. In our example, the entropy is maximal and is equal to $\ln 4 = 1.386...$ We built such a frame of the "H" sign type (Fig. 5). The font developed on the basis of such a sample could have the name **SuperCompl*[6], that is, extremely equi-dimensional.

An object consisting of a single part ($p = 1$) has a minimum equi-dimensionality equal to 0. For frame analysis, when equi-dimensionality is minimal, this corresponds to an empty FontWindow or one that is uniformly filled with one part (component) (examples, Figs. 3a, 3b, 3c, 3d). In our case, if a free field is occupied by a single part, then, *like many other ideals, this ideal of simplicity makes the use of such a sign type senseless*. The "H" frame of the *Super C* font (*Appendix*, No. 64), which is almost completely represented by free space, is the closest to it. On the verge of such a situation there are the so-called “conceptual fonts” that use the rejection of some letter elements, for example, "H" of the *Gropius Display* font (*Appendix*, No. 10), whose "H" frame does not have a crossbar and free space between the boles, and *GHSans* (*Appendix* No. 22), whose "H" is devoid of a crossbar. The first one has the rank formula $O_S S_S B_S = C_S$ and entropy $\boldsymbol{H} = 0.665$; for the second one, $O_S S_S C_S B_S$ and $\boldsymbol{H} = 0.866$.

Anentropy A

The third integral characterization of the composition, that is, anentropy, was proposed by T.G. Petrov [23] to reduce the degree of uncertainty in the description of the composition of a multicomponent system when it is mapped by two characterizations, that is, the rank formula and entropy. It is clear that the more components in the system and the fewer its characterizations, the greater the undetermined information about a particular object, system, organism, etc., is.

Anentropy is calculated by the formula

$$A = -1/n \sum \ln p_i - \ln n,$$

where the designations are the same as in the entropy formula. The active contribution of the p_i component to the anentropy is “$-\ln p_i$.”

Anentropy, determined by the arithmetic mean of the active contribution taken with a negative sign, is relatively weakly dependent on large p_i and strongly dependent on small p_i. It changes from 0, with a uniform distribution, i.e., when the entropy is maximal, up to an indefinitely large quantity, when at least one p_i approaches zero. For fonts, the latter situation is very rare, but, in order not to deny their ordering, all characterizations of their character types are given in the Appendix. For the diagrams, they can be placed in them above the entropy value, close to the upper edge of the diagram, breaking the frame line somewhat, to mark the anomaly of the point and focus on the fact that as p_i tends to zero, the entropy tends to $+\infty$. As an example, this refers to the anomalous sign types of the fonts *Gropius Display* No. 10 and *GHSans* No. 22, which almost cease to be letters, or to the figures shown in Fig. 3, where (a) the shares of parts of B, C and O areas approach 0; (b) the S_S, B_S, O_S areas are disappearing; (c) the S_S, C_S, O_S areas approach zero; (d) the S_S, B_S, C_S areas tend to zero.

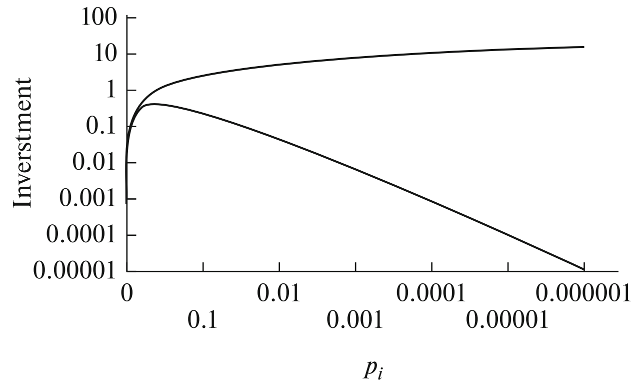


Fig. 4. The dependencies of active contribution to entropy “$-p_i \ln p_i$”, (lower curve) and to anentropy “$-\ln p_i$” (upper curve) from shares of the areas p_i.

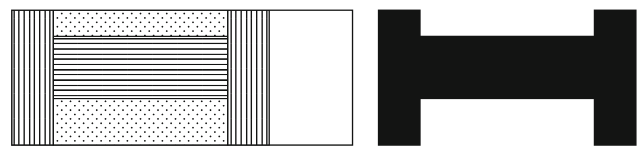


Fig. 5. H components under the condition of equality of all areas in the FontWindow.

Thus, anentropy is calculated for the same areas in the FontWindow as entropy, but, unlike it, the negative logarithm of the area share is used as an contribution, i.e., “$-\ln p_i$.” This value approaches 0 when p_i approaches 1 (this is for all foundations of logarithms), i.e., to the area of the entire FontWindow, and increases monotonically with decreasing p_i.

THE METHOD OF ORDERING *RHA* CODE-WORDS OF H SIGN TYPES

The combination of the rank formula, entropy, and anentropy is considered as the FontWindow code with

[6] The “*” sign is placed before a linguistic (semiotic) object that is non-existent, but was invented as a mental experiment.

Table 1. All possible rank formulas—*R*-classes of "H" sign types ordered according to the SBCO alphabet

1	$\mathbf{S_sB_sC_sO_s}$	7	$B_sS_sC_sO_s$	13	$C_sS_sB_sO_s$	19	$\mathbf{O_sS_sB_sC_s}$
2	$S_sB_sO_sC_s$	8	$B_sS_sO_sC_s$	14	$C_sS_sO_sB_s$	20	$\mathbf{O_sS_sC_sB_s}$
3	$S_sC_sB_sO_s$	9	$B_sC_sS_sO_s$	15	$C_sB_sS_sO_s$	21	$O_sB_sS_sC_s$
4	$S_sC_sO_sB_s$	10	$B_sC_sO_sS_s$	16	$C_sB_sO_sS_s$	22	$O_sB_sC_sS_s$
5	$S_sO_sB_sC_s$	11	$B_sO_sS_sC_s$	17	$\mathbf{C_sO_sS_sB_s}$	23	$\mathbf{O_sC_sS_sB_s}$
6	$\mathbf{S_sO_sC_sB_s}$	12	$B_sO_sC_sS_s$	18	$C_sO_sB_sS_s$	24	$O_sC_sB_sS_s$

its contents and, at the same time, as a word in the information *RHA* language. Three alphabets are used to build this word.

<u>The first alphabet</u> is the order of the characters of the components given by the authors of the article: S C B O given in the notation of Fig. 2. We adopted this sequence as the alphabet for the ordering of the rank formulas for the frames of letters. First, the entire set of specific *RHA* code-words of the FontWindow (accompanied by font names) is divided into groups with identical first members of rank formulas, after which the groups are arranged according to the alphabet (SBCO), as when using alphabets of existing scripts. Then, within each group, the operation is repeated with the second member of the rank formula, and so on. As a result, for the four-letter (rating) part of the codes, we have a complete set of possible rank formulas, which are arranged in the table as words in the alphabetical sequence by SBCO, where rank formulas for which there are examples, are in bold.

<u>The second alphabet</u> is continual, that is, the entropy values. Within a group of identical rank formulas for the alphabetical ordering of sign types, entropy is used, which is expressed as a real number. However, the direction of ordering descriptions of objects within a group with the same rank formulas, in principle, may be different. Thus, in the article on alphabets [32], the arrangement of codes by non-decreasing of *H* was used, as in the case of streamlining the codes of the age distribution of the population [26]. In [32], the direction of ordering was taken without any justification, in [26] the direction was justified using historical and demographic reasons. With the accumulation of material on H, and the analysis of the *HA* diagrams, it became clear that the variant of ordering fonts used in [32] makes it difficult to interpret the results. Therefore, it was decided to alphabetically order the *H* values by their lack of an increase[7], which we explain below.

<u>The third alphabet</u> is also continual, that is, the anentropy values. To order codes with similar *R* and *H*, the *A* values are used according to their lack of decrease, which was chosen due to the fact that the inverse relationship between *H* and *A* dominates statistically.

[7] This means that different fonts may have equal entropy values.

RHA-ORDERING OF H SIGN TYPES

Codes of 99 arbitrarily selected Latin fonts from the ParaType, MyFonts, FontShop collections are presented in the *Appendix*. Rank formulas are ordered by the SBCO alphabet. Entropies within the same rank formula, that is, the *R*-class of the font, are ordered by the lack of an increase. In the “Frame in F” column, the frames are inserted into the FontWindow so that the left border of the FontWindow coincides with their left border. This allows us to draw a number of conclusions.

(1) Out of the total number, 24, of possible *R*-classes of H (see Table 1), 6 classes are presented (including the first one, proposed only in this article).

(2) Approximately 90% of all the H sign types considered by the authors in this article belong to two *R*-classes:

$O_sS_sC_sB_s$ (Nos. 11–64) external space, boles, connecting bar, internal space,

$O_sC_sS_sB_s$ (Nos. 65–99) external space, internal space, boles, connecting bar.

The first class are predominant, in which there are 54 fonts with relatively thick boles, and the second, in which there are 35 “leaner” ones.

(3) The theoretical construction of the H frame, which we called *SuperCompl, with the rank formula $S_sB_sC_sO_s$, with a sequence of characters matching the sequence of characters in the alphabet, the maximum possible (for four components) entropy and the minimum possible anentropy, defines the role of this frame as the beginning for ordering fonts according to rank formulas (see Table 1), entropy (see the *Appendix*), and anentropy. The sequences of characters at the beginning of words in alphabetic dictionaries are most similar to sequences of characters at the beginning of the alphabet, ideally they should be the alphabet itself. Accordingly, the last *R*-component of the *RHA* word (rank formula) will represent the sequence of characters inverse to the alphabet.

(4) The location of the theoretical H frame of **SuperCompl* as the first in the ordering is determined by the method of ordering the rows of the table and within the *R*-classes. Its exclusivity follows from the fact that: (a) this sign contains all components that are clearly visually distinguishable and (b) this image, which has the maximum entropy, with any changes in proportions between the four components, will be transformed into any different sign, which will have

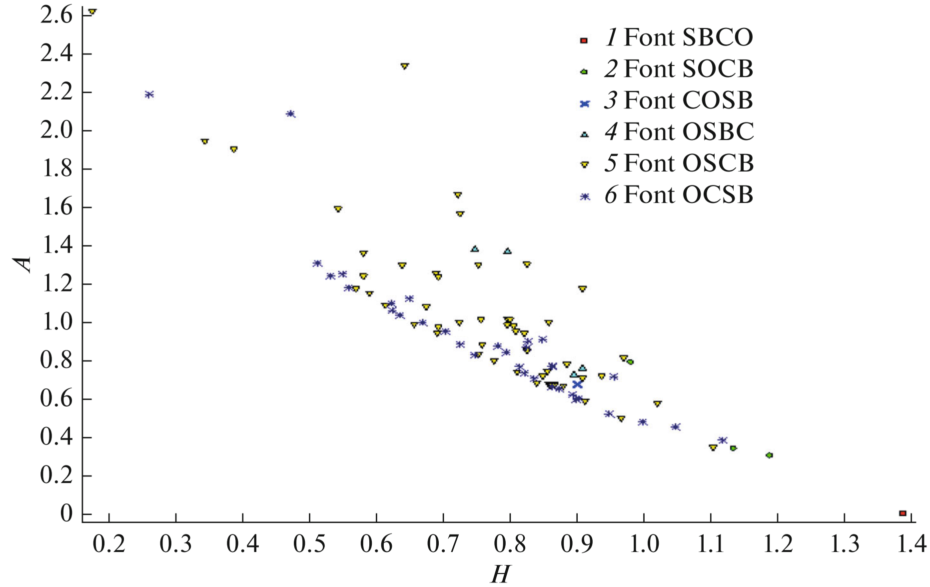


Fig. 6. Entropy–anentropic characterization of the frameworks given in the *Appendix*.

less entropy, as in the $O_sS_sB_sC$ class, to which it is assigned in the catalog layout, and in all others, arising from the initial equality of the areas of the FontWindow components.

Let us again note that when the share of the area of any component tends to 1, the frame will experience the same effect as when filling the FontWindow with one tone, i.e., the entropy will decrease to zero and the anentropy will grow indefinitely. The same will happen when the areas of any (one or more) components tend to 0.

(5) in sequences of the most widely represented classes of frameworks $O_sS_sC_sB_s$ and $O_sC_sS_sB_s$ by the lack of an increasing of entropy (against the background of a general decrease in the visually perceptible width of the frame, up to the frame of the fonts *Super C* No. 64 and *FR Pasta Mono* No. 99, which are practically indistinguishable visually), the frames of the first class have a much greater variety of entropy parameters than those of the second class. This follows from both the *Appendix* and the representation of the distribution of the corresponding points in Fig. 6.

(6) Fonts that are identical or most similar in appearance mostly have the same rank formulas and the same or similar entropic characterizations. In the *Appendix*, they are located adjacently, such as: *PT_PTS75 Bold* No. 25 and *PT_PTS65 Demi* No. 26 (with codes: *R,* $O_sS_sC_sB_s$, $H = 0.860$, $A = 0.664$ and $H = 0.858$, $A = 0.671$, respectively), *Tahoma* No. 78 and *Kabel* No. 79 ($O_sC_sS_sB_s$, 0.837, 0.702 and $O_sC_sS_sB_s$, 0.835, 0.704), or close, for example, Regata No. 21 and PT_PTS95 Black No. 30 ($O_sS_sC_sB_s$, 0.866, 0.667 and 0.848, 0.713).

Exceptions, for example, the greatest closeness of letter forms in the fonts 49 and 86, or 52 and 89, are due to the fact that when the areas of different components of the frame are similar, small changes in one of the areas can lead to a rearrangement in the rank formula.

(7) Visually, the sequence of codes for the change in H in the two most represented R-groups in the table reveal some general tendencies in changing the shape of the frame. This commonality is due to the fact that in the FontWindows of these groups the share of free space prevails and increases; therefore in both cases the frames become increasingly narrow by the end of these groups. In cases where other components of the FontWindow are in the first place in the rank formula, an increase in their share of the area leads to the formation of monstrous frames, such as those shown in Figs. 3a, 3b, 3c.

(8) Against the background of these trends the frames stand out that are sharply different from those belonging to general trends to one degree or another. This is a consequence of the fact that the framework has three characterizations (R, H, A,) provided that $\sum p_i = 1$, and under linear ordering of descriptions, only one is used. Deviations from the tendency in changing the forms of frames in the *Appendix*, especially pronounced in the $O_sS_sC_sB_s$ block, are a reflection of the inadequacy of the representation of frame structures that have more than two components, with a single information entropy. As examples of frames whose entropy does not differ or differs by no more than 0.001 while the entropy differs significantly we give Nos. 17–18, 31–32, 34–35, 43–44, 46–47, and 48–49. When the values of entropy and anentropy are similar, the differences between the frames become insignifi-

cant and almost disappears: Nos. 24–25, 58–59, 74–75, and 92–93. In this case, we have a vivid visual illustration that a boundless breadth of diversity that is estimated by a single number may occur when the variety of values is very significant.

The question of deviations from the tendencies of changes in the frames at the linear ordering of their entropy is largely eliminated by using the entropy–anentropic diagrams given below.

FONT FRAMEWORKS IN THE ENTROPY–ANENTROPY *HA*-DIAGRAM

For a general review of the entropy–anentropic characterization of the font collection, we present Fig. 6.

As we can see, the points on the diagram are distributed unevenly, with a relatively small area in the intervals of H 0.6–1.0 and A 0.6–1.2.

The **SuperCompl* frame, which was built theoretically on the principle of maximizing H and minimizing A and which is still the only representative of the R-class $S_sB_sC_sO$, is in the lower right corner of the diagram. Fonts whose entropy–anentropic characterizations are located to the right and below it in the proposed diagram do not exist and cannot exist for a frame with four components. At the same time, another no less important feature of this frame is that insignificant changes in the areas of its components lead to the generation of all 5 existing and 18 under-represented classes in the Appendix. No other class has been found.

Since the transition to **SuperCompl* from any existing frame class can occur due to small changes in any component, deviations from **SuperCompl* can continue in an undefined direction, generating new frames. A similar situation occurs in the case of strange attractors [35] or bifurcations [36, 37] as described in functional analysis. The **SuperCompl* frame acts as a *pure ideal; loss of completeness of it leads to the appearance of all the other 23 classes* of frames, each of which has some deviation from the ideal **SuperCompl* due to kenosis (a decrease) of at least one of the components and the elevation of the others. Thus, all possible frames are in one ideal variant and are its emanations [38]. The same relationship exists between a vector with module 0, which is not directed anywhere, and an arbitrarily small vector that has one and only one directionality, while all other directionalities are excluded, the same as a geometric point with respect to any other shape.

The maximum values of entropy and at the same time the closest to the theoretical value of the **SuperCompl* font are the frames of the two R-class fonts $S_sO_sC_sB_s$ and one each of the classes $O_sS_sC_sB_s$ and $O_sC_sS_sB_s$. The ratios of the areas of the four FontWindow components of such fonts are closest to each other and to the ratio of the theoretical **SuperCompl*.

The minimum entropy values have frames with the $O_sS_sC_sB_s$ and $O_sC_sS_sB_s$ rank formulas, they are the narrowest and most recent ones in the font lists of the corresponding R-classes. The free space sharply dominates in these fonts in the FontWindow.

The "H" frames of conceptual fonts have the highest anentropy values ($A = +\infty$): *Gropius Display* (No. 10) has only two components of font coding (boles and free space of the FontWindow) and *GHSans* (No. 22), has three components. These are missing from the diagram.

The R-classes $O_sS_sC_sB_s$ and $O_sC_sS_sB_s$ are the most represented in the system, while the fields of their entropy–anentropic characterizations overlap to a large extent, they have differences. With equal values of entropy, the "H" frames of $O_sS_sC_sB_s$ in most cases have higher values of anentropy (Fig. 6).

The high density of points on the diagram did not allow showing the position of the points with images of allographs and their frames to make the picture more visual without significantly reducing the number of points on the diagram. Figure 7 shows HA and images of allographs of R-class $O_sS_sC_sB$; Fig. 8 shows images of frames of the same R-class.

When comparing the characteristics of H and A of the "H" frames in Fig. 8, it is possible to observe general trends in the shape of frames and their grouping.

Frames whose components are commensurate in area in the FontWindow occupy the area of high entropy values. Accordingly, frames whose components have very small shares among other components have large anentropy values. Two of the variants found in the literature: *GHSans* (No. 22), which is devoid of the crossbar and *Gropius Display* (No. 10), which is devoid of two components, have the maximum anentropy value of $+\infty$. Therefore, they are not placed on the diagram.

Along the bottom edge of the dotted area (Fig. 7, Fig. 8) from right to left, from the bottom up, a tendency is clearly observed from the widest to the narrowest fonts, which correlates with an increase in the area of free space in the FontWindow.

In the case of equality or closeness of R, H, and A, the compositions and images of the H sign types are either similar or close in terms of the ratio of the frame areas, although they differ aesthetically.

The codes of the clone fonts have the same values as the “honest” fonts; therefore the clones in the diagram will be at the same point or almost coincide with the previous “honest” ones.

Thus, during the transition from the allograph to its frame, there is a change in the salience of sign types: in the case of allographs, letter forms that have features (Nos. 28, 36, 44, 69, 89, 92, and 95) and missing components (Nos. 10 and 22) stand out, and in the case of frames, they have extreme entropy characterizations (No. 1 Max$H$min$A$ and No. 64, min$H$Max$A$)

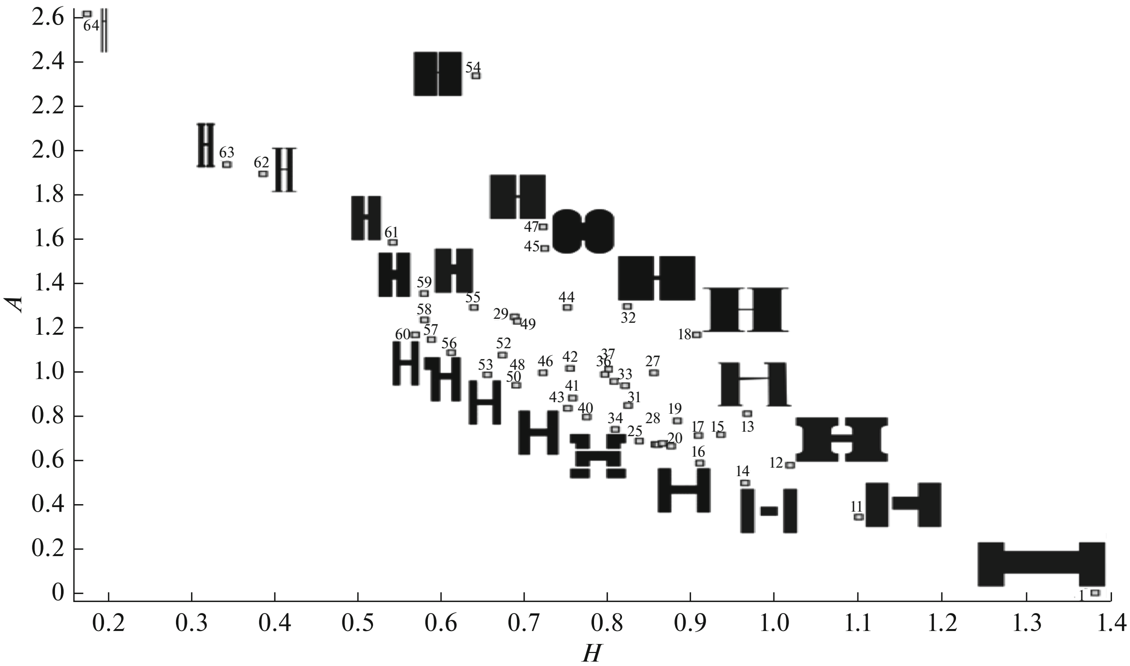


Fig. 7. H allographs with the rank formula $O_SS_SC_SB_S$ and S_S=B_S=C_S=O_S (No. 1, in the lower right corner) on the entropy–anentropic diagram (the numbers at the points correspond to the numbers in the *Appendix*).

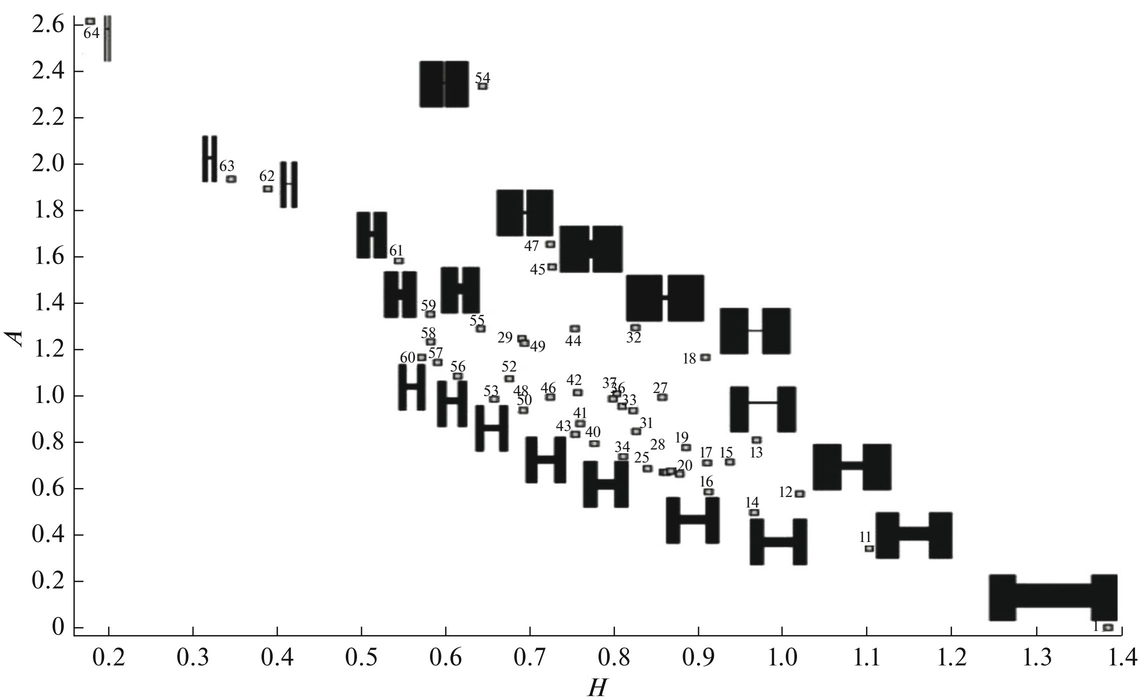


Fig. 8. H frames with the rank formula $O_SS_SC_SB_S$ and S_S=B_S=C_S=O_S (No. 1) the entropy–anentropic diagram.

The schematization of the coding of the component composition of the frames leads to a sharp reduction in the qualitative diversity; in particular, this is due to the exclusion of the extreme (sharply standing out or "pathological") variants of allographs.

DISCUSSION OF RESULTS

The presented collection ("collection" in the sense of Foucault [39]) has 99 fonts, that is, it is a representative of a universum with an unknown structure and limits of variation. It is still small for general conclusions; thus, for now, only some preliminary considerations can be made.

The rank formulas we present in the table exhaust their variety with 24 classes, of which with an arbitrary choice of 99 fonts presented in the *Appendix*, only 5 classes are implemented (the 6th one we designed), while the overwhelming number of them are in only two classes. Therefore, we see that only a small part of the logically admissible variety of roman fonts is used.

The large problem was choosing the direction of ordering entropy and anentropy due to the lack of obvious indisputable reasons for determining the sign of changes, to decrease or to increase (for rocks this direction was determined by the fact that crystallization of magmas occurs when the temperature decreases, and this, in turn, directs the composition of the magmatic solution to its separation and a statistically significant decrease in the entropy of the new systems that are minerals compared with the initial silicate solution).

Fonts as works of art are generally not related to each other and their authors apparently did not know how their design evolved over time; thus, until the theoretical form of the letter and the accumulation of a "sufficiently" complete collection of sign types was created, the problem of their ordering was not even formulated. Initially borrowed from other cases, as data was accumulated the ordering by increasing entropy (as in [32]) yielded to a more heuristic type under the pressure of another and higher coherence of the whole when a new reference point was detected. The "H" frame of **SuperCompl* was a starting point, occupying the lowest and extreme right position in the diagram (Fig. 6). This frame acts as basic string of generative grammer [40], which is initial point for generating of all of the other frames

The small variety of rank formulas (it seems to increase slightly with the intended accumulation of materials) indicates the rather narrow information-aesthetic requirements for fonts of consumers and designers who provide services.

When varying the ratio of the frame areas, it becomes possible to predict and build the shape of the H frames that are located between adjacent points on the ***HA*** diagram or on any line tied to this diagram, which can be considered a frame family line.

Thus, even the available material shows that the ***RHA*** method provides an alternative method for making new judgments about existing and potential fonts.

CONCLUSIONS

The question arises of why all this is necessary if there are already ways to search for similar font patterns. This can be given a general answer.

The conversion of science to art is a means to understand "how IT is done," "why SUCH a letter form is better," "why THIS WAY is more convincing, more beautiful, more emotional, more appropriate ...." Is it possible to understand, and it is interesting, why less than one-third of the 24 classes of fonts are used? What will happen if one composes layouts of letters and frames for classes that are still empty? Perhaps these layouts will give an impulse to the artist's imagination to create something acceptable and new?

As an example of this, the frame of the *Arial* font (see the *Appendix*, No. 73) was chosen and, by adopting a constant set of its metric characterizations (17.56, 2.33, 10.09, 70.02), we permuted their values for the components of the new frames. Thus, two frames with new rank formulas, were built, namely, *COBS*  and *SBOC* , as well as one with the same rank formula as the only still theoretical (**SuperCompl*) *SBCO* . If the first two seem to be of little promise, then the last variant, as we believe, can be considered quite decent for use in some significant texts of headings—inscriptions.

There are no clear correspondences between characterizations of letters, complete "words" in the *RHA* language, and usually distinguished properties of letters; however, islands of local correspondences between the images of letters and codes exist; this is the advantage of a meaningful code. This also applies to cases when one needs to find a font with quantitatively defined properties, for example, when perceiving a moving text, or a fixed text from a vehicle moving with a known characteristic speed; when perceiving texts that have to be read at an angle that is significantly different from direct, written on non-planar surfaces, observed with curved mirrors when inspecting hard-to-reach parts of structures, etc. In connection with this, we can mention the problem of studying the semiotics of moving inscriptions, which turned out to be one of the essential problems in the study of the linguistic landscape of a city[8] (the observer's position, perspective point, including being on a moving vehicle, as one of the four main images, principal imaging systems, in cognitive linguistics [41, p. 254–255]).

The proposed system for encoding H of direct lettering when organizing a single data bank will allow one to:

(1) find the frames that are similar or analogous in lettering;

(2) purposefully obtain monotonously varying series of frames, thereby expanding the variety of fonts;

(3) begin the study of the psychology of text perception at a quantitative level in the following areas:

(a) gathering statistics and identifying visual-thematic connections, for which types of texts with letter forms are used more often;

[8] See, for example, a series of articles in a special issue of the International Journal Bilingualism, 2014, Vol. 18 (5).

(b) tracking trends and the development of fashion for certain letter forms;

(c) highlighting the zones of the most readable letter forms for predicting the creation of new fonts for specific purposes;

(d) identifying letter forms that are not used in design at all and trying to determine the reasons for this.

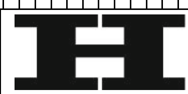
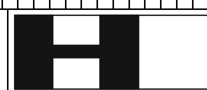
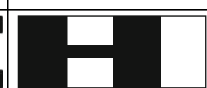
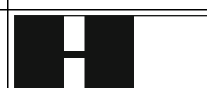
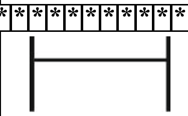
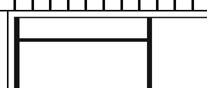
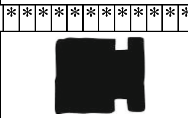
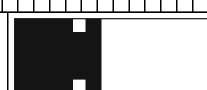
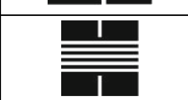
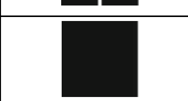
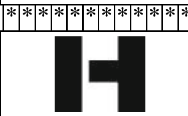
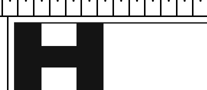
The integrality of font characterizations used in *RHA* codes certainly requires some skill to work successfully with them, as with all new tools.

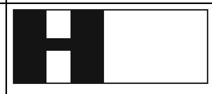
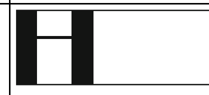
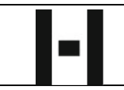
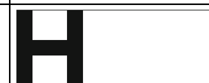
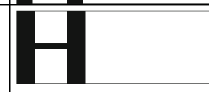
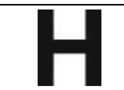
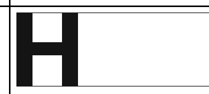
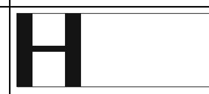
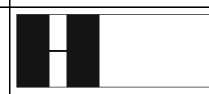
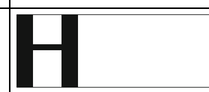
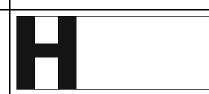
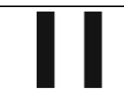
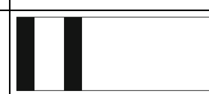
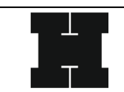
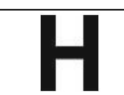
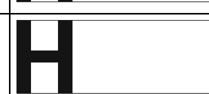
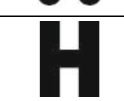
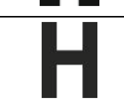
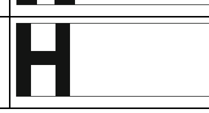
The method considered in this article can be used in coding geographic maps, where the rank formula is built, for example, according to the ratio of the areas of the intervals between the contours of the map, according to the ratio of the areas of different landscapes on the terrain, or geological maps, for which the ratio of areas of images of rocks of different compositions, ages, or varying degrees of mineralization is significant. It is possible to transfer to coding volumes or mass ratios in an object, for example, a particle size analysis of rocks. This method is also suitable for describing the coloring of stones [42], animal skins, interiors, the leaf mosaic of trees, the structure of a rash in dermatology, etc.

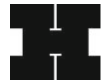
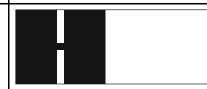
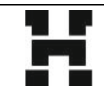
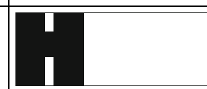
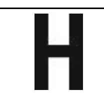
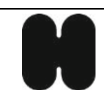
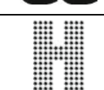
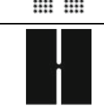
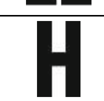
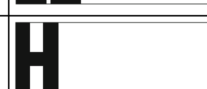
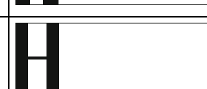
Most of the work on the *RHA* method was carried out using the Petros3.2 program compiled by S.V. Moshkin.

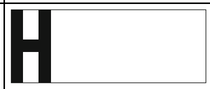
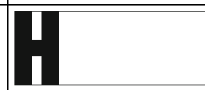
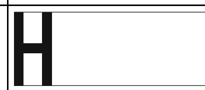
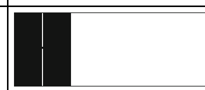
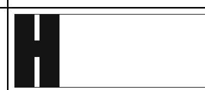
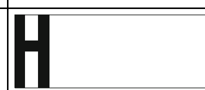
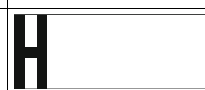
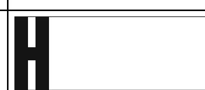
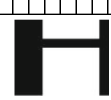
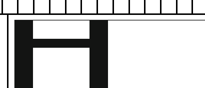
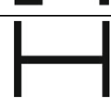
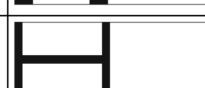
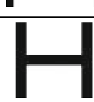
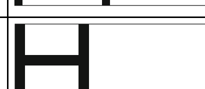
APPENDIX

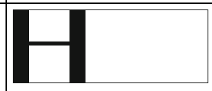
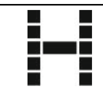
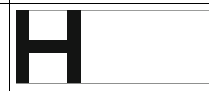
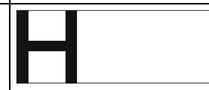
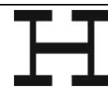
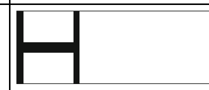
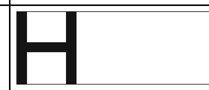
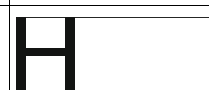
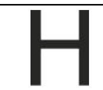
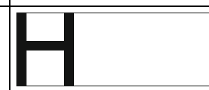
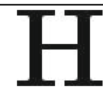
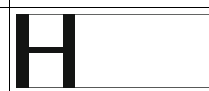
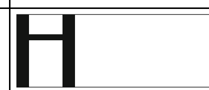
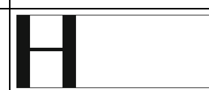
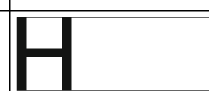
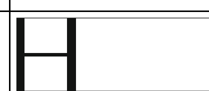
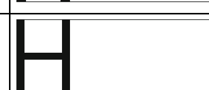
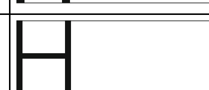
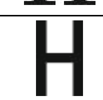
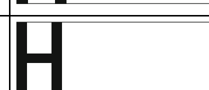
The *RHA*-systematics of H frames

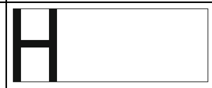
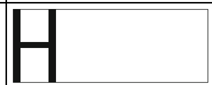
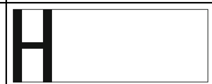
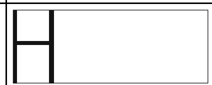
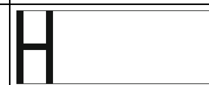
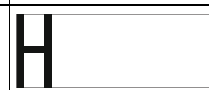
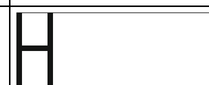
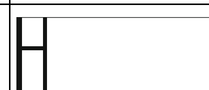
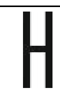
No.	*R*-class	*H*	*A*	Font	Allograph	Frame in F	S_S %	B_S %	C_S %	O_S %
1	S_s=B_s=C_s=O_s	1.386	0.000	*SuperCompl l			25.00	25.00	25.00	25.00
			**							
2	S_s=$O_sC_sB_s$	1.187	0.303	King Tut Black			39.42	4.01	20.25	36.32
3	$S_sO_sC_sB_s$	1.134	0.340	Blackoak Std			52.01	4.01	20.24	23.74
4	$S_sO_sC_sB_s$	0.980	0.787	Guinness Extra Stout NF			50.7	0.88	9.73	38.7
			**							
5	$C_sO_sS_sB_s$	0.900	0.678	Vienna Extended LET			4.56	3.07	62.8	29.6
			**							
6	$O_sS_sB_sC_s$	0.908	0.759	Cuadrifonte			38.31	4.03	2.19	55.47
7	$O_sS_sB_s$=C_s	0.896	0.729	Tonal			33.60	3.47	3.03	59.89
8	$O_sS_sB_sC_s$	0.796	1.373	YWFT Pudge			44.77	2.06	0.33	52.85
9	$O_sS_sB_s$=C_s	0.747	1.378	Sinaloa			36.87	0.85	0.82	61.46
10	O_sS_s=B_sC_s	0.665	+∞	Gropius Display			38.20	0.00	0.00	61.80
			**							
11	$O_sS_s\mathit{C_sB_s}$	1.103	0.341	Aksent			27.47	5.22	12.80	54.51

No.	*R*-class	*H*	*A*	Font	Allograph	Frame in F	S_S %	B_S %	C_S %	O_S %
12	$O_sS_sC_sB_s$	1.020	0.571	Egyptienne Extd D Bold			34.19	2.09	10.47	53.26
13	$O_sS_sC_sB_s$	0.970	0.808	Quadrata			21.63	0.69	17.00	60.68
14	$O_sS_sC_sB_s$	0.966	0.491	Glaser Stencil			16.34	3.63	14.00	66.03
15	$O_sS_sC_sB_s$	0.937	0.713	Georgia			19.16	1.21	15.05	64.58
16	$O_sS_sC_sB_s$	0.911	0.582	PT_PTC75 Caption Bold			16.31	2.76	12.31	68.62
17	OS=CB	0.909	0.704	Book Antiqua			15.86	1.41	15.59	67.14
18	$O_sS_sC_sB_s$	0.908	1.169	Falstaff			33.66	0.22	8.54	57.58
19	$O_sS_sC_sB_s$	0.884	0.773	School-Book_Bold			16.74	1.12	13.88	68.25
20	$O_sS_sC_sB_s$	0.878	0.660	Project Fairfax			18.84	2.21	9.65	69.31
21	$O_sS_sC_sB_s$	0.866	0.667	Regata			19.61	2.40	8.25	69.74
22	$O_sS_sC_sB_s$	0.866	$+\infty$	GHSans			18.20	0.00	15.16	66.65
23	$O_sS_sC_sB_s$	0.864	0.669	Thickset			23.33	3.42	4.94	68.32
24	$O_sS_sC_sB_s$	0.861	0.677	PT_PTS85 Extra Bold			17.75	2.31	9.35	70.58
25	$O_sS_sC_sB_s$	0.860	0.664	PT_PTS75 Bold			15.07	2.24	11.43	71.27
26	$O_sS_sC_sB_s$	0.858	0.671	PT_PTS65 Demi			15.06	2.17	11.46	71.31
27	$O_sS_sC_sB_s$	0.857	0.993	BodoniBold			18.16	0.46	12.85	68.53
28	$O_sS_sC_sB_s$	0.856	0.736	Yess_Bold			24.35	2.15	5.81	67.69
29	$O_sS_sC_sB_s$	0.848	0.713	PT_PTS95 Black			20.67	2.18	7.16	69.99
30	$O_sS_sC_sB_s$	0.839	0.677	Century Gothic			14.80	2.35	10.34	72.51

No.	*R*-class	*H*	*A*	Font	Allograph	Frame in F	S_S %	B_S %	C_S %	O_S %
31	$O_sS_sC_sB_s$	0.825	0.845	YWFT Black Slabbath			27.25	1.80	4.05	66.89
32	$O_sS_sC_sB_s$	0.825	1.294	Braggadocio			42.77	0.30	3.20	53.73
33	$O_sS_sC_sB_s$	0.822	0.935	Acsioma_Shock			31.78	1.50	3.06	63.65
34	$O_sS_sC_sB_s$	0.810	0.734	FF Archian Stencil Pro			16.54	2.13	8.04	73.29
35	$O_sS_sC_sB_s$	0.809	0.947	Acsioma_Next Rough			30.52	1.53	2.91	65.05
36	$O_sS_sC_sB_s$	0.806	0.975	Acsioma_Me-dium			31.47	1.44	2.71	64.38
37	$O_sS_sC_sB_s$	0.800	1.007	Acsioma_Super Shock			31.75	1.29	2.64	64.33
38	$O_sS_sC_sB_s$	0.797	0.980	School-Book_Cond Bold			31.78	1.33	2.50	64.38
39	$O_sS_sC_sB_s$	0.797	1.012	Acsioma_Next			16.94	0.63	10.04	72.39
40	$O_sS_sC_sB_s$	0.777	0.793	PT_PTN87 Narrow Extra Bold			15.92	1.84	7.48	74.76
41	$O_sS_sC_sB_s$	0.758	0.876	PT_PTN97 Narrow Black			19.07	1.59	5.22	74.12
42	$O_sS_sC_sB_s$	0.756	1.008	Adamant			23.24	1.01	4.13	71.62
43	$O_sS_sC_sB_s$	0.753	0.828	PT_PTN77 Narrow Bold			13.57	1.63	8.44	76.36
44	$O_sS_sC_sB_s$	0.752	1.289	Avatar			34.82	0.80	1.28	63.10
45	$O_sS_sC_sB_s$	0.726	1.558	Fatta			36.97	0.46	0.73	61.84
46	$O_sS_sC_sB_s$	0.724	0.991	Diamonds			19.03	1.10	4.71	75.16
47	$O_sS_sC_sB_s$	0.723	1.657	Gaslon			31.29	0.12	2.07	66.52
48	$O_sS_sC_sB_s$	0.694	0.970	PT_PTS87 Cond Extra Bold			15.10	1.22	5.62	78.07
49	$O_sS_sC_sB_s$	0.693	1.227	BodoniCondC			12.69	0.32	9.11	77.88

No.	*R*-class	*H*	*A*	Font	Allograph	Frame in F	S_S %	B_S %	C_S %	O_S %
50	$O_sS_sC_sB_s$	0.691	0.935	PT_PTS77 Cond Bold			12.34	1.31	7.26	79.10
51	$O_sS_sC_sB_s$	0.689	1.249	Garbage			17.65	0.36	5.42	76.56
52	$O_sS_sC_sB_s$	0.675	1.074	PT_PTS97 Cond Black			17.76	1.03	3.75	77.46
53	$O_sS_sC_sB_s$	0.657	0.983	PT_PTS67 Cond Demi			9.68	1.18	8.30	80.84
54	$O_sS_sC_sB_s$	0.642	2.333	Loudine			27.98	0.02	0.87	71.12
55	$O_sS_sC_sB_s$	0.640	1.291	Impact			19.94	0.64	2.27	77.14
56	$O_sS_sC_sB_s$	0.614	1.082	Hill			11.08	0.99	5.72	82.21
57	$O_sS_sC_sB_s$	0.591	1.144	PT_PTS79 Extra Cond Bold			10.91	0.85	5.22	83.02
58	$O_sS_sC_sB_s$	0.582	1.235	PT_PTS89 Extra Cond Extra Bold			13.51	0.74	3.40	82.35
59	$O_sS_sC_sB_s$	0.581	1.352	PT_PTS99 Extra Cond Black			16.26	0.61	2.18	80.96
60	$O_sS_sC_sB_s$	0.570	1.165	PT_PTS69 Extra Cond Demi			8.31	0.79	6.68	84.22
61	$O_sS_sC_sB_s$	0.544	1.585	CompactBold			15.33	0.26	2.10	82.31
62	$O_sS_sC_sB_s$	0.389	1.895	Radar			5.82	0.10	3.79	90.29
63	$O_sS_sC_sB_s$	0.344	1.938	Titanic Condensed			6.22	0.14	2.11	91.52
64	$O_sS_sC_sB_s$	0.177	2.613	Super C			1.96	0.04	1.49	96.51
				***						
65	$O_sC_sS_sB_s$	1.118	0.382	Flat10 ArtDeco			19.20	3.34	25.40	52.05
66	$O_sC_sS_sB_s$	1.047	0.451	Distill			7.84	4.39	36.50	51.27
67	$O_sC_sS_sB_s$	0.998	0.476	Ecyr			10.46	3.77	23.83	61.94

No.	*R*-class	*H*	*A*	Font	Allograph	Frame in F	S_S %	B_S %	C_S %	O_S %
68	$O_sC_sS_sB_s$	0.955	0.713	Ustav			16.08	1.13	19.70	63.09
69	$O_sC_sS_sB_s$	0.947	0.520	YWFT LED			13.26	3.34	16.45	66.94
70	$O_sC_sS_sB_s$	0.903	0.600	Yess_Regular			11.86	2.61	16.64	68.89
71	$O_sC_sS_sB_s$	0.899	0.593	Lodge			6.40	3.79	22.26	67.54
72	$O_sC_sS_sB_s$	0.893	0.618	PT_PTC55 Caption Regular			11.04	2.50	17.25	69.20
73	$O_sC_sS_sB_s$	0.874	0.651	Arial			10.09	2.33	17.56	70.02
74	$O_sC_sS_sB_s$	0.864	0.661	Verdana			10.16	2.30	16.82	70.72
75	$O_sC_sS_sB_s$	0.863	0.768	SchoolBook_-Book			12.45	1.26	16.51	69.77
76	$O_sC_sS_sB_s$	0.862	0.765	Bengaly			12.82	1.28	15.97	69.93
77	$O_sC_sS_sB_s$	0.848	0.908	Bodoni			13.17	0.69	16.32	69.82
78	$O_sC_sS_sB_s$	0.837	0.702	Tahoma			10.18	2.07	15.49	72.26
79	$O_sC_sS_sB_s$	0.835	0.704	Kabel			9.28	2.12	16.50	72.10
80	$O_sC_sS_sB_s$	0.826	0.898	Pacioli			8.20	0.89	21.09	69.82
81	$O_sC_sS_sB_s$	0.824	0.864	Times New Roman			10.81	0.95	16.77	71.46
82	$O_sC_sS_sB_s$	0.822	0.732	PT_PTS55 Regular			9.34	1.92	16.01	72.73
83	$O_sC_sS_sB_s$	0.814	0.766	PT_PTS45 Light			8.31	1.74	17.36	72.59
84	$O_sC_sS_sB_s$	0.795	0.841	Calipso			6.16	1.50	20.29	72.05
85	$O_sC_sS_sB_s$	0.781	0.876	School-Book_Cond			11.33	1.06	13.15	74.46
86	$O_sC_s{=}S_sB_s$	0.748	0.825	PT_PTN67 Narrow Demi			10.61	1.62	10.88	76.89

No.	*R*-class	*H*	*A*	Font	Allograph	Frame in F	S_S %	B_S %	C_S %	O_S %
87	$O_sC_sS_sB_s$	0.726	0.883	PT_PTN57 Narrow Regular			7.97	1.41	13.11	77.51
88	$O_sC_sS_sB_s$	0.703	0.950	PT_PTN47 Narrow Light			7.39	1.14	13.24	78.23
89	$O_sC_sS_sB_s$	0.669	0.999	Moon Star Soul			8.76	1.01	10.12	80.11
90	$O_sC_sS_sB_s$	0.650	1.118	Jatran			4.28	0.84	15.68	79.21
91	$O_sC_sS_sB_s$	0.635	1.035	PT_PTS57 Cond Regular			6.80	1.06	10.59	81.56
92	$O_sC_sS_sB_s$	0.625	1.059	Circus			7.16	0.98	9.83	82.03
93	$O_sC_sS_sB_s$	0.624	1.097	PT_PTS47 Cond Light			5.47	0.90	12.11	81.52
94	$O_sC_sS_sB_s$	0.560	1.176	Zirkus			6.55	0.80	7.99	84.65
95	$O_sC_sS_sB_s$	0.551	1.249	Rustica			4.65	0.66	10.20	84.49
96	$O_sC_s{=}S_sB_s$	0.533	1.240	PT_PTS59 Extra Cond Regular			6.75	0.69	6.87	85.69
97	$O_sC_sS_sB_s$	0.512	1.304	PT_PTS49 Extra Cond Light			5.47	0.59	7.61	86.34
98	$O_sC_sS_sB_s$	0.473	2.083	FOSU			0.61	0.12	15.27	84.00
99	$O_sC_sS_sB_s$	0.263	2.185	FR Pasta Mono			1.99	0.09	3.71	94.22

ACKNOWLEDGMENTS

The authors thank K.M. Kirichenko for compiling a program for analyzing images of letters; N.A. Pavlova, T.M. Zhuravskaya, and K.A. Manukyan for help with terminology.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Fontproblem. http://www.promotion.su/link/25403. Accessed December 21, 2018.
2. ParaType. http://www.paratype.ru/e-zine/issue01/synonims.htm.
3. Korol'kova, A., *Zhivaya tipografika* (Living Typography), Moscow: IndexMarket, 2011.
4. Bringhurst, R., *The Elements of Typographic Style,* Hartley & Marks, 1992.
5. Kapr, A., *Schriftkunst,* Dresden: VEB Verlag der Kunst, 1971.
6. YouWorkForThem. Section Fonts. Subsection Serif. http://www.youworkforthem.com/fonts/serif. Accessed December 21, 2018.
7. YouWorkForThem. Section Fonts. Subsection Sans-serif. http://www.youworkforthem.com/fonts/sans-serif. Accessed December 21, 2018.
8. YouWorkForThem. Section Fonts. Subsection Stencil. http://www.youworkforthem.com/fonts/stencil. Accessed December 21, 2018.

9. YouWorkForThem. Section Fonts. Subsection Handwriting. http://www.youworkforthem.com/fonts/handwriting. Accessed December 21, 2018.
10. MyFonts. Section Typographic Category. http://www.myfonts.com/category/. Accessed December 21, 2018.
11. The Letterhead Studio Fonts Catalog 2001–2002. http://issuu.com/letterhead/docs/lh_font_catalog. Accessed December 21, 2018.
12. The Letterhead Studio Catalog 1999. http://issuu.com/letterhead/docs/lhs_catalog_1999. Accessed December 21, 2018.
13. *Fotonabornye shrifty* (Phototypesetting Fonts), Kozubov, G. and Efimov, V., Eds., Moscow: Kniga, 1983.
14. *Tsifrovye shrifty* (Digital Fonts), Moscow: ParaTaip, 2004.
15. Monotype fonts by A-Z. http://catalog.monotype.com/alphabetical. Accessed October 10, 2019.
16. Abstract Fonts section A-Z. http://www.abstractfonts.com/alpha/A. Accessed February 20, 2019.
17. Font Spring. Section Alphabetical. http://www.fontspring.com/a2z?sort=alpha. Accessed February 20, 2019.
18. Habr: How to find the right font without knowing its name? https://habr.com/en/post/38046/. Accessed February 20, 2019.
19. My fonts by Monotype. http://www.myfonts.com/WhatTheFont/. Accessed February 20, 2019.
20. Identifont. http://www.identifont.com/identify.html. Accessed February 20, 2019.
21. Shmeleva, A., Classification of fonts: Practice and problems, *Publish,* 2003, no. 1.
22. Moles, A., *Theorie de l'information et perception esthetique,* Paris: Flammarion, 1958.
23. Petrov, T.G., Justifying a variant of the general classification of geochemical systems, *Vestn. Leningr. Gos. Univ.,* 1971, no. 18, pp. 30–38.
24. Petrov, T.G. and Farafonova, O.I., *Informatsionno-komponentnyi analiz. Metod RHA* (Information Component Analysis. The RHA Method), St. Petersburg, 2005.
25. Petrov, T.G., The RHA data language for describing the compositions of multicomponent systems, *Autom. Doc. Math. Linguist.,* 2001, vol. 35, no. 2, pp. 8–22.
26. Chebanov, S.V. and Petrov, T.G., Intensionality, intensional alphabets, and intensional words and dictionaries, in *Aktual'nye problemy sovremennoi kognitivnoi nauki* (Current Problems of Modern Cognitive Science), Ivanovo, 2013, pp. 239–266. https://doi.org/10.13140/RG.2.1.4542.8644
27. Petrov, T.G., The RHA method for coding, systematizing, and displaying changes in the age composition of the population. https://www.researchgate.net/publication/282136556_Metod_RHA_dla_kodirovania_sistematizacii_i_otobrazenia_izmenenij_vozrastnyh_sostavov_naselenia. Accessed February 20, 2019. https://doi.org/10.13140/RG.2.1.3207.2166
28. Petrov, T.G., Graphic representation of the evolutionary processes of the compositions of multicomponent objects of any nature, *Autom. Doc. Math. Linguist.,* 2012, vol. 46, no. 2, pp. 79–93. https://doi.org/10.3103/S0005105512020045
29. Petrov, T.G. and Moshkin, S.V., RHA(T)-system for coding of discrete distributions and their alteration processes, *Proc. 3rd International Multi-Conference on Complexity, Informatics and Cybernetics IMCIC 2012,* Orlando, FL, 2012, pp. 12–16.
30. Petrov, T.G. and Krasnova, N.I., *R-slovar'-katalog khimicheskikh sostavov mineralov* (The R-Dictionary-Directory of Chemical Compositions of Minerals), St. Petersburg: Nauka, 2010.
31. Petrov, T.G., "Soft" system of coordinates in regular simplexes, *Int. J. Intell. Inf. Syst.,* 2015, vol. 4, no. 1, pp. 1–7. http://article.sciencepublishinggroup.com/pdf/10.11648.j.ijiis.20150401.11.pdf. https://doi.org/10.11648/j.ijiis.20150401.11
32. Petrova, E.T. and Petrov, T.G., Encoding and systematization of fonts based on the RHA method information language, *Vestn. S.-Peterb. Gos. Univ., Tekhnol. Diz., Ser 2,* 2015, no. 1, pp. 39–44.
33. Sedov, E.A., *Odna formula i ves' mir. Kniga ob entropii* (One Formula and the Whole World. The Book on Entropy), Moscow: Znanie, 1982.
34. Petrov, T.G., On the degree of complexity of geochemical systems from the view of information theory, *Dokl. Akad. Nauk SSSR,* 1970, vol. 191, no. 4, pp. 1094–1096.
35. Thom, R., *Stabilité structurelle et morphogenèse,* Avalon Publishing, 1994.
36. Arnol'd, V.I., Afraimovich, V.S., Il'yashenko, Yu.S., and Shil'nikov, L.P., Bifurcation theory, in *Sovremennye problemy matematiki. Fundamental'nye napravleniya* (Current Problems of Mathematics. Fundamental Directions), Moscow, 1985, vol. 5, pp. 5–218.
37. Arnol'd, V.I., Catastrophe theory, in *Sovremennye problemy matematiki. Fundamental'nye napravleniya* (Current Problems of Mathematics. Fundamental Directions), Moscow, 1985, vol. 5, pp. 219–277.
38. Dumitrake, Yu., Kenosis of Christ in the understanding of Sergei Bulgakov and Dumitru Stăniloae, *Diskussiya,* 2013, no. 3, pp. 14–18.
39. Foucault, M., *Les mots et les choses une archeologie des sciences humaines,* Paris: Gallimard, 1966.
40. Chomsky, N., *Syntactic Structures,* Walter de Gruyter, 1957.
41. Talmy, L., How language structures space, in *Spatial Orientation: Theory, Research, and Application,* Pick, H.L., Jr. and Acredolo, L.P., Eds., New York–London, 1983, pp. 225–282.
42. Petrov, T.G. and Shuiskii, A.V., Parametric description of the colored stone pattern, *Sovrem. Probl. Nauki Obraz.,* 2013, no. 5. https://science-education.ru/ru/article/view?id=10235.

Translated by S. Avodkova