MINISTRY OF EDUCATION, CULTURE AND RESEARCH OF THE REPUBLIC OF MOLDOVA
STATE UNIVERSITY "DIMITRIE CANTEMIR"

Manuscript
UDC: 512.536.7+515.122.4,004.424.62

BUDANAEV IVAN

DISTANCES ON FREE MONOIDS AND THEIR APPLICATIONS IN THEORY OF INFORMATION

111.03 MATHEMATICAL LOGIC, ALGEBRA AND NUMBER THEORY

Summary of Ph.D. Thesis in Mathematics

Thesis was drafted and submitted at Mathematics and Information Science Doctoral School.

CHIŞINĂU, 2019

MINISTERUL EDUCAȚIEI, CULTURII ŞI CERCETĂRII
UNIVERSITATEA DE STAT "DIMITRIE CANTEMIR"

Cu titlul de manuscris
C.Z.U.: 512.536.7+515.122.4,004.424.62

BUDANAEV IVAN

DISTANȚE PE MONOIZI LIBERI ŞI APLICAȚIILE LOR ÎN TEORIA INFORMAȚIEI

111.03 LOGICA MATEMATICĂ, ALGEBRA ŞI TEORIA NUMERELOR

Rezumatul tezei de doctor în științe matematice

CHIŞINĂU, 2019

The thesis was elaborated in the Doctoral School of Mathematics and Information Science, "Dimitrie Cantemir" University, in consortium with the Institute of Mathematics and Computer Science "V. A. Andrunachievici", "Alecu Russo" State University in Balti, State University of Tiraspol and "Bogdan Petriceicu Hasdeu" State University in Cahul.

Doctoral advisor:

CHOBAN Mitrofan, academician, Dr. habil., professor.

Composition of the Doctorate Committee:

COJOCARU Svetlana, Dr. habil., Professor, Corresponding Member of the Academy of Sciences of Moldova - **chairman**.

ARNAUTOV Vladimir, academician, Dr. habil., professor.

KENDEROV Petar, academician of Bulgarian Academy of Science, Dr. habil., professor.

SÎRBU Parascovia, Ph. D., associate professor.

CIOBAN Mitrofan, academician, Dr. habil., professor.

The public thesis presentation will take place on 14.05.2019, 14:00, at the meeting of the Doctoral Committee, at the "Dimitrie Cantemir" University, multipurpose hall, Hincesti 54 Street, Chisinau.

Doctoral thesis and summary can be accessed at the Central Scientific Library "Andrew Lupan" and on ANACEC web page (http://www.anacip.md).

Doctoral summary was sent out on 12.04.2019.

Secretary, Naval Elvira

Ph. D., associate professor.

Doctoral advisor, Cioban Mitrofan

academician, Dr. habil., professor.

Author Budanaev Ivan

Contents

KEYWORDS

Keywords: Alexandroff space, quasivariety of topological monoids, free monoids, invariant distance, quasi-metric, Levenshtein distance, Hamming distance, Graev distance, parallel decomposition, proper similarity, weighted mean, bisector of two strings, convexity, algorithm.

1. RESEARCH GOALS AND OBJECTIVES

Actuality and importance of the research topic. The notion of universal algebra has been introduced in the book of Alfred North Whitehead "A Treatise on Universal Algebra", published in 1898 [45]. Between 1935 and 1950, Birkhoff introduced varieties and quasivarities of universal algebras, free algebras, universal algebra congruence, subalgebra lattice, homomorphism theorems. Due to the second world war, the results published by Anatol Maltsev in the years 1938 – 1946 were not noted until the early 50s of the last century. Alfred Tarski's plenary lecture in 1950 at the Cambridge International Mathematics Congress inaugurated a new era. After 1950, various aspects of model theory were studied, with uncommon applications in mathematical logic, language theory, automata theory, with contribution of the following mathematicians: A. Robinson, A. Tarski, G. Birkhoff, C.C. Chang, L. Henkin, S. C. Kleene, B. Jonsson, A. Church, S. Eilenberg, S. MacLane, R. Lyndon, A. I. Maltsev, V. I. Arnautov, M. A. Arbib, V. M. Gluşkov, N. Chomsky, M. Minsky, S. Ginsburg, D. Scott, D. A. Huffman, E. Marczewski, J. Mycielski, P. J. Higgins, B. I. Plotkin, Yu. I. Manin, S. Marcus, A. G. Kurosh, V. I. Glivenko, V. D. Belousov, A. P. Ershov, O. B. Lupanov, A. D. Wallace and others (see [3, 8, 6, 10, 17, 20, 21, 24, 28, 29, 32, 34, 44]). Various applications of algebra in information analysis and image processing were explored by S. Cojocaru, C. Gaindric, V. Shcherbacov, P. Syrbu, V. Izbash and others [18, 23, 40, 43]. In solving many problems related to information analysis (processing, digitization, comparison, classification), it became necessary to study invariant metrics and topologies on free universal algebras. The study of topological algebras was initiated with the study of Lie groups, topological groups and topological linear spaces.

Different types of distances were examined by M. Frechet, V. Niemytzcki, P.S. Alexandroff, A.V. Arhangelskii, M.M. Choban, R.W. Heath, P. Kenderov, S. Nedev, W.A. Wilson (see [22, 2, 5, 27, 4, 15, 36]). In the class of distances, quasi-metrics are highlighted by the fact that they are not symmetric but satisfy the condition of the axiom of the triangle inequality. Discrete quasi-metrics bring us to the concept of digital space and more general to Alexandroff space. Hamming [26], Graev [25] and Levenshtein [31] research work bring us to the need to develop methods of extension of distances on the alphabet A over the free monoid $L(A)$. We mention that it is important for the extension to be invariant. This problem is important and remained unsolved until now for any quasivariety of topological monoids.

These facts determine the actuality and importance of the research topic.

The thesis presents theoretical results of the study of distances on abstract algebraic structures. The applicative part of the research can be used in information theory, where it is necessary to define the measure similarity between data and the efficiency of data representation. These notions, in their turn, can be obtained by applying distance between the information sequences.

The research goals and objectives. The goal of the scientific research is to study the problem of distances on free monoids. Some of the goals posed are to study the properties of these distances on the free monoid $L(A)$, determine if there exists any relationship between them and to find applications of the obtained results. To achieve this goal, the following objectives were defined:

- determine the conditions of extension of given quasi-metric ρ on A and any quasivariety $\mathcal{V}$ of topological monoids to an invariant quasi-metric ρ^* on the free monoid $F^a(A, \mathcal{V})$;
- elaboration of an effective method for extending the quasi-metric ρ on free monoids;
- determine the conditions of the existence of invariant topologies on free monoid $F^a(A, \mathcal{V})$ which are extensions of the given topology on A;
- establish relations between Hamming, Levenshtein and Graev distances;
- development of efficient representations and decompositions of pairs of strings;
- implementation of innovative algorithms for solving text sequences problems;
- implementation of algorithms for construction of weighted means and bisector sets of a given pair of strings;
- describe image processing from the topological point of view;
- analyze properties of digital topologies on the discrete line.

The goals set in the beginning of the research, once achieved, leads to important theoretical results. This way, the notion of the parallel decompositions of a pair of strings was introduced. Based on that, the extended research goals covered the study of the measure of proper similarity of a pair of strings, the construction algorithms of the weighted means and bisector sets of pairs of strings, as well as the examination and solution to the question of the convexity property of the mentioned sets.

2. SCIENTIFIC RESEARCH METHODOLOGY

The present study is conducted within the area of the algebraic and topological theories, and the methodology used is based on the application of methods of monoids theory, distance spaces, language theory, algorithms theory and the informational systems theory. Using the scientific research methodology, the work was partitioned into the following stages: problem identification, hypothesis formulation, hypothesis investigation and analysis, conclusions of the established results.

Hypothesis formulation is based on the goals and objectives of the research. One of the main hypothesis, which influenced the discovery of the results following from it, was the existence of the extension of a quasi-metric ρ on free monoids. This way, some of the early stage research results comprise of effective methods of distance extension on free monoids, which lead to the possibility of introducing the concept of parallel decomposition of strings.

Hypothesis investigation and analysis were conducted within the scope of the fundamental scientific research methodology. During this stage it was established that for any non-Burnside quasivariety $\mathcal{V}$ and any quasi-metric ρ on a set X with basepoint p_X on free monoid $F^a(X, \mathcal{V})$

there exists a unique stable quasi-metric $\hat{\rho}$ with the Properties 2.4.1 through 2.4.10, which are summarized as a general result in Theorem 2.4.1.

The thesis scientific novelty and originality consist in its new theoretical results which are published in peer-reviewed scientific journals. Research results comprise of effective methods of distance extension on free monoids, which lead to the possibility of introducing the concept of parallel decomposition of strings. This has allowed the development of the concepts of efficiency and similarity of the information sequences, as well as the construction of the sets of weighted mean and bisector of strings. The degree of the novelty and originality is represented by:

- method of quasi-metric extension on free monoid $F^a(X, \mathcal{V})$;
- study of the digital and Alexandroff spaces;
- presented solutions for Maltsev problems for quasivariaties of topological monoids;
- established relations between Hamming, Graev and Levenshtein distances on free monoids;
- introduction of the concept of efficiency of representation;
- introduction of the concept of the optimal parallel decompositions for strings from the free monoids;
- algorithms implemented for weighted mean and bisector construction for pairs of strings;
- proof of the non-convexity of the informational segment;
- introduction of the notion of the symmetric topology on the digital line;
- proof of the uniqueness of Khalimsky topology as minimal digital topology;
- elaboration of the digital image processing algorithm from the topological perspective, applicable in the digital space.

The important scientific problem solved in the research is the development of methods for constructing and studying distances extension over free monoids, which contribute to obtaining effective methods of information representation, applicable to solving different distance problems such as sequence alignment, proper similarity of a pair of strings, construction of weighted means and bisectors of a pair of strings.

The theoretical significance is determined by obtaining new results regarding the establishment of the conditions of existence of the extension of distances on free monoids, that permit the construction of distinct invariant topologies on free monoids. The elaborated methods have allowed to approach the problems related to information sequences from a new point of view. Additionally, the theoretical results permit the study of the digital line, and the minimality property of Khalimsky topology.

The applicative value of the paper consists in the use of the obtained theoretical results in the study of symmetric topologies on the digital line, imaging processing and construction of the centroid of a set of strings. Presented methods build larger sets of elements, using the method of optimal parallel decompositons.

Approval of scientific results. The scientific results obtained were presented at national and international scientific conferences, and were published in peer-reviewed journals. The main results included in the thesis were presented at the following conferences scientific:

- *Scattered and Digital Topologies in Information Sciences.* Plenary talk at the Conference of the Romanian Society of Applied and Industrial Mathematics ROMAI, CAIM 2018, Chisinau, Moldova, 20-23 September 2018;
- *Scattered and Digital Topologies in Image Processing.* Conference on Mathematical Foundations of Informatics, MFOI 2018, Chisinau, Moldova, 2-6 July 2018;
- *About Non-Convexity of the Weighted Mean of a Pair of Strings.* International Conference "Contemporary Trends in Science Development: Visions of Young Researchers", Academy of Sciences of Moldova, Chisinau, Moldova, 15 June 2018;
- *On the Midset of Pairs of Strings.* International Conference on Mathematics, Informatics and Information Technologies, MITI 2018, Balti, Moldova, 19-21 April 2018;
- *Measures of Similarity on Monoids of Strings.* Conference on Mathematical Foundations of Informatics, MFOI 2017, Chisinau, Moldova, 9-11 Nov 2017;
- *Parallel Decompositions of Pairs of Strings and Their Applications.* Conference on Applied and Industrial Mathematics, Iasi, Romania, 14-17 Sept 2017;
- *On the Bisector of a Pair of Strings.* The 4th Conference of Mathematical Society of the Republic of Moldova, dedicated to the centenary of Vladimir Andrunachievici (1917-1997) CMSM4, Chisinau, Moldova, 28 June - 2 July 2017;
- *Distances on Monoids of Strings and Their Applications.* Conference on Mathematical Foundations of Informatics, MFOI 2016, Chisinau, Moldova, 25-31 July 2016;
- *Invariant Distances on Free Semigroups and Their Applications.* The 20th Annual Conference of the Mathematical Sciences Society of Romania, 19-22 May 2016;

Publications on the topic of thesis research. The results obtained in the thesis are published in 20 papers scientific articles (see [46]–[65]): 7 articles in journals (see [47, 49, 56, 60, 62, 63, 64]), 13 papers in international conferences (see [46, 48, 50, 51, 52, 54, 53, 55, 57, 58, 59, 61, 65]); 8 publications by single author (see [46]–[53]), including 2 articles in peer-reviewed journals (see [47, 49]). The total volume of publications is 6.4 sheets of author.

Thesis structure and volume: the thesis is written in English and consists of: introduction, four chapters, general conclusions and recommendations, 192 bibliography titles. The total volume of the thesis is 128 pages, out of which 122 main text pages.

3. SYNTHESIS OF CHAPTERS

In the introduction, the actuality and importance of the research topic are formulated. In addition, the research goals, objectives, the scientific novelty and originality are stated. The scientific problem under study is presented with the emphasis on the importance of the theoretical and applicative value of the work. A brief analysis of the problems and publications on the thesis topic is given. This sections concludes with a summary of the content of the paper.

The first chapter, Current situation in the field of quasi-metric space theory and their applications in algebra and information theory, has an introductory character and contains a survey of the most important results related to the purpose and objectives of the Thesis, the directions of the investigation. It defines and classifies distances, distance spaces, informational systems of Scott-Ershov type, universal topological algebras, spaces of strings.

For solving stated problem it is important to study the following particular problems:

1. To determine the conditions of extension of given quasi-metric ρ on A and any quasivariety $\mathcal{V}$ to a invariant quasi-metric ρ^* on the free monoid $F^a(A, \mathcal{V})$.
2. To propose the algorithms of the calculation of the distance $\rho^*(a, b)$ between two information sequences $a, b \in L(A)$.
3. To determine the relations between topologo-geometrical properties of spaces (A, ρ) and $(L(A), \rho^*)$.
4. To propose methods of construction of weighted means and bisector sets of a given pair of strings.
5. To determine topologo-geometrical properties which are important in the analysis of information and image processing.

Maltsev problems are formulated for free monoids, which were posed in 1957 [33]:

First Maltsev's Problem: Under which conditions the mapping v_X is an embedding?

Second Maltsev's Problem: Under which conditions the homomorphism w_X is a continuous isomorphism?

For complete regular spaces X the Maltsev Problems were solved affirmatively by S. Swierczkowski [42] in the case of discrete signature E, and by M. M. Choban and S. S. Dumitrashcu for any signature [19, 14].

At the end of this chapter, the research problem is formulated, the methods of solving it are identified, the goals and objectives of the research are established.

In second chapter, Extension of quasi-metrics on free topological monoids, quasivarieties of topological monoids are studied and new methods of quasi-metrics extension are elaborated. Chapter begins with the introduction into free topological monoids and construction of the abstract free monoid.

Any topological space X is considered to be a Kolmogorov space, i.e. a T_0-space: for any two distinct points $x, y \in X$ there exists an open subset U of X such that $U \cap \{x, y\}$ is a singleton set. A class $\mathcal{V}$ of topological monoids is called a quasivariety of monoids if: (1) the class $\mathcal{V}$ is multiplicative; (2) if $G \in \mathcal{V}$ and A is a submonoid of G, then $A \in \mathcal{V}$; (3) every space $G \in \mathcal{V}$ is a T_0-space. A free monoid of a space X in a class $\mathcal{V}$ is a topological monoid $F(X, \mathcal{V})$ with the properties: $X \subseteq F(X, \mathcal{V}) \in \mathcal{V}$ and p_X is the unity of $F(X, \mathcal{V})$, the set X generates the monoid $F(X, \mathcal{V})$, and for any continuous mapping $f : X \longrightarrow G \in \mathcal{V}$, where $f(p_X) = e$, there exists a unique continuous homomorphism $\bar{f} : F(X, \mathcal{V}) \longrightarrow G$ such that $f = \bar{f}|X$. An abstract free monoid of a space X in a class $\mathcal{V}$ is a topological monoid $F^a(X, \mathcal{V})$ with the properties: X is a subset of $F^a(X, \mathcal{V})$, $F^a(X, \mathcal{V}) \in \mathcal{V}$ and p_X is the unity of $F^a(X, \mathcal{V})$, the set X generates the monoid $F^a(X, \mathcal{V})$, and for any mapping $f : X \longrightarrow G \in \mathcal{V}$, where $f(p_X) = e$, there exists a unique continuous homomorphism $\hat{f} : F^a(X, \mathcal{V}) \longrightarrow G$ such that $f = \hat{f}|X$.

By the Kakutani method it is proved that:

Theorem 2.1.1. *Let $\mathcal{V}$ be a non-trivial quasivariety of topological monoids. Then for each space X the following assertions are equivalent:*

1. *There exists $G \in \mathcal{V}$ such that X is a subspace of G and p_X is the neutral element in G.*
2. *For the space X there exists the unique free topological monoid $F(X, \mathcal{V})$.*

We continue by highlighting the problems addressed in this section.

Problem 2.1.1. *Let $\mathcal{V}$ be a non-trivial quasivariety of topological monoids. Under which conditions for a space X there exists the free topological monoid $F(X, \mathcal{V})$?*

Fix a space X for which there exists the free topological monoid $F(X, \mathcal{V})$. Then there exists a unique continuous homomorphism $\pi_X : F^a(X, \mathcal{V}) \longrightarrow F(X, \mathcal{V})$ such that $\pi_X(x) = x$ for each $x \in X$. The monoid $F(X, \mathcal{V})$ is called abstract free if π_X is a continuous isomorphism.

Problem 2.1.2. *Let $\mathcal{V}$ be a non-trivial quasivariety of topological monoids. Under which conditions for a space X there exists the free topological monoid $F(X, \mathcal{V})$, which is abstract free?*

The Problems 2.1.1 and 2.1.2 are important in the theory of universal algebras with topologies (see [33, 14, 15, 12, 13, 16]). These problems for varieties of topological algebras were posed by A. I. Maltsev ([33], see Maltsev's problems in thesis section 1.6).

The following theorem's proof relies on the result stated in previous theorem 2.1.1.

Theorem 2.1.2. *Let $\mathcal{V}$ be a non-trivial quasivariety of topological monoids and there exists $H \in \mathcal{V}$ and point $b \in H$ such that $e \neq b$, and $E = \{e, b\}$ is a discrete subspace of H. Then for each zero-dimensional space X there exists the unique free topological monoid $F(X, \mathcal{V})$.*

Example 2.1.1 illustrates that not for any non-trivial quasivariety $\mathcal{V}$ and any T_0-space X there exists $F(X, \mathcal{V})$.

Let $\omega = \{0, 1, 2, \ldots\}$. A quasivariety $\mathcal{V}$ of topological monoids is called a Burnside quasivariety if there exist two minimal numbers $p = p(\mathcal{V}), q = q(\mathcal{V}) \in \omega$ such that $0 \leq q < p$ and $x^p = x^q$ for all $x, y \in G \in \mathcal{V}$. In this case, any $G \in \mathcal{V}$ is a (p, q)-periodic monoid of the exponent (p, q). If $q = 0$, then any monoid $G \in \mathcal{V}$ is a periodic monoid of the exponent p and $x^p = e$ for each $x \in G \in \mathcal{V}$. The trivial quasivariety is considered Burnside of the exponent $(0, 1)$.

The following theorem solves Problem 2.1.1 for complete non-Burnside quasivarieties of topological monoids.

Theorem 2.3.3. *Let $\mathcal{V}$ be a complete non-Burnside quasivariety of topological monoids. Then for each T_0-space X there exists the free topological monoid $F(X, \mathcal{V})$.*

The following theorem solves Problem 2.1.1 for complete non-trivial quasivarieties of topological monoids.

Theorem 2.3.4. *Let $\mathcal{V}$ be a complete non-trivial quasivariety of topological monoids. Then for each completely regular space X there exists the free topological monoid $F(X, \mathcal{V})$.*

Fix a non-trivial complete quasivariety $\mathcal{V}$ of topological monoids. Consider a non-empty set X with a fixed point $e \in X$. We assume that $e \in X \subseteq F^a(X, \mathcal{V})$ and e is the identity of the monoid $F^a(X, \mathcal{V})$. Let ρ be a pseudo-quasi-metric on the set X. Denote by $Q(\rho)$ the set of all stable pseudo-quasi-metrics d on $F^a(X, \mathcal{V})$ for which $d(x, y) \leq \rho(x, y)$ for all $x, y \in X$. The set $Q(\rho)$ is non-empty, since it contains the trivial pseudo-quasi-metric $d(x, y) = 0$ for all $x, y \in F^a(X, \mathcal{V})$. For all $a, b \in F^a(X, \mathcal{V})$ we put $\hat{\rho}(a, b) = sup\{d(a, b) : d \in Q(\rho)\}$. We say that $\hat{\rho}$ is the maximal stable extension of ρ on $F^a(X, \mathcal{V})$. For any $a, b \in F^a(X, \mathcal{V})$ we put $\bar{\rho} = inf\{\Sigma\{\rho(x_i, y_i) : i \leq n\} : n \in \mathbb{N}, x_1, y_1, x_2, y_2, ..., x_n, y_n \in X, a = x_1x_2...x_n, b = y_1y_2...y_n\}$ and $\rho^*(a, b) = inf\{\bar{\rho}(a, z_1) + ... + \bar{\rho}(z_i, z_{i+1}) + ... + \bar{\rho}(z_n, b) : n \in \mathbb{N}, z_1, z_2, ..., z_n \in F^a(X, \mathcal{V})\}$.

One of the main results of the thesis is the following theorem.

Theorem 2.4.1. *Let ρ be a pseudo-quasi-metric on X, Y be a subspace of X and $e \in Y$. Denote by $M(Y) = F^a(Y, \mathcal{V})$ the submonoid of the monoid $F^a(X, \mathcal{V})$ generated by the set Y and by d_Y the extension of $\hat{\rho}|Y$ on $M(Y)$ of the pseudo-quasi-metric ρ_Y on Y, where $\rho_Y(y, z) = \rho(y, z)$ for all $y, z \in Y$. Then:*

1. *$d_Y(a, b) = \hat{\rho}(a, b)$ for all $a, b \in M(Y)$.*
2. *If $\mathcal{V}$ is a non-Burnside quasivariety, then $\bar{\rho}(x, y) = \rho(x, y)$ for all $x, y \in X$.*
3. *If ρ is a (strong) quasi-metric on Y, then $\hat{\rho}$ is a (strong) quasi-metric on $M(Y)$.*
4. *If ρ is a metric on Y, then $\hat{\rho}$ is a metric on $M(Y)$.*
5. *If $a, b \in F^a(Y, \mathcal{V})$ are distinct points and ρ is a quasi-metric on $Sup(a, b)$, then $\hat{\rho}(a, b) + \hat{\rho}(b, a) > 0$.*
6. *If $a, b \in F^a(Y, \mathcal{V})$ are distinct points and ρ is a strong quasi-metric on $Sup(a, b)$, then $\hat{\rho}(a, b) > 0$ and $\hat{\rho}(b, a) > 0$.*

7. *For any $a, b \in F^a(Y, \mathcal{V})$ there exist $n \in \mathbb{N}$, $x_1, x_2, ..., x_n \in Sup(a, a)$ and $y_1, y_2, ..., y_n \in Sup(b, b)$ such that $a = x_1x_2...x_n$, $b = y_1y_2...y_n$, $n \leq l(a) + l(b)$ and $\bar{\rho}(a, b) = \Sigma\{\rho(x_i, y_i) : i \leq n\}$.*
8. *$\hat{\rho} = \bar{\rho} = \rho^*$.*

In the class of free monoids, theorem 2.4.1 permits to solve the first and second Maltsev problems, which were formulated for universal algebras in 1957 [33].

The quasivariety of topological monoids $\mathcal{V}$ is rigid if for any space X, any word $a \in F(X, \mathcal{V})$, any point $c \in X \setminus \{p_x\}$ and any representation $ac = x_1x_2...x_n$, where $x_1, x_2, ..., x_n \in X$, there exists $m \leq n$ such that $x_m = c$ and $a = x_1x_2...x_{m-1}$. In this case $x_i = p_X = e$ for each $i > m$.

The variety of all topological monoids is rigid.

Theorem 2.5.1. *Let $\mathcal{V}$ be a non-Burnside rigid quasivariety of topological monoids, ρ be a quasi-metric on a space X with basepoint p_X and $\rho(x, p_X) = \rho(y, p_X)$ for all $x, y \in X \setminus \{p_X\}$, or $\rho(p_X, x) = \rho(p_X, y)$ for all $x, y \in X \setminus \{p_X\}$. Then $\rho^*(ac, bc) = \rho^*(ca, cb) = \rho^*(a, b)$ for all $a, b, c \in F(X, \mathcal{V})$.*

The following theorem improves Theorem 2.3.3 and solves Problem 2.1.2 for complete non-Burnside quasivarieties of topological monoids.

Theorem 2.6.1. *Let $\mathcal{V}$ be a non-trivial complete non-Burnside quasivariety of topological monoids. Then:*

1. For each T_0-space X on the free monoid $F^a(X, \mathcal{V})$ there exists a T_0-topology $\mathcal{T}(qm)$ such that:

– $(F^a(X, \mathcal{V}), \mathcal{T}(qm)) \in \mathcal{V}$;

– X is a subspace of the space $(F^a(X, \mathcal{V}), \mathcal{T}(qm))$;

– the topology $\mathcal{T}(qm)$ is generated by the family of all invariant pseudo-quasi-metrics on $F^a(X, \mathcal{V})$ which are continuous on X.

2. For each T_0-space X the free topological monoid $F(X, \mathcal{V})$ exists and is abstract free.

3. A space X is a T_1-space if and only if spaces $F(X, \mathcal{V})$ and $(F^a(X, \mathcal{V}), \mathcal{T}(qm))$ are T_1-spaces.

4. A space X is functionally Hausdorff if and only if the spaces $F(X, \mathcal{V})$ and $(F^a(X, \mathcal{V}), \mathcal{T}(qm))$ are functionally Hausdorff.

Results analogous to theorems 2.1.1 and 2.6.1 are obtained for semi-topological monoid $F(X, \mathcal{W})$ in theorems 2.7.1 and 2.7.2.

The second chapter ends with the discussion on topological digital spaces, and the results summarized in Corollary 2.8.2, which follow from Corollary 2.8.1 and Propositions 2.8.2 and 2.8.4.

Corollary 2.8.1. *Let $\mathcal{V}$ be a non-trivial complete non-Burnside quasivariety of topological monoids. Then for each space X the following assertions are equivalent:*

1. $F(X, \mathcal{V})$ is an Alexandroff space.

2. On a space $F(X, \mathcal{V})$ there exists a quasi-metric with the natural values.

3. X is an Alexandroff space.

Corollary 2.8.2. *Let $\mathcal{V}$ be a non-trivial complete non-Burnside quasivariety of topological monoids. Then for each space X the following assertions are equivalent:*

1. $F(X, \mathcal{V})$ is a topological digital space.

2. X is a topological digital space.

The results presented in this chapter successfully complement the works of mathematicians in the domain of the distance extension on the abstract algebraic structures. The author's work in this chapter is published in the articles [57, 62] and serve as a base for research presented in the next chapters.

In chapter 3, Measures of similarity on monoids of strings, it is proved that there are invariant distances on the monoid $L(A)$ of all strings closely related to Levenshtein and Hamming distances. A distinct definition of the distance on $L(A)$ is introduced, based on the Markov-Graev method, proposed for free groups. In result, it is shown that for any quasi-metric d on alphabet A in union with the empty string there exists a maximal invariant extension d^* on the free monoid $L(A)$. This new approach allows to introduce parallel and semiparallel decompositions of two strings, which are later used in subsection 3.3 to determine the efficiency and penalty factors in two strings representations.

For any strings $a, b \in L(A)$ we find the decompositions of the form $a = v_1u_1v_2u_2 \dots v_ku_kv_{k+1}$ and $b = w_1u_1w_2u_2 \dots w_ku_kw_{k+1}$, which can be represented as $a = a_1a_2 \dots a_n$, $b = b_1b_2 \dots b_n$ with the following properties:

- some a_i and b_j may be empty strings, i.e. $a_i = \varepsilon$, $b_j = \varepsilon$;
- if $a_i = \varepsilon$, then $b_i \neq \varepsilon$, and if $b_j = \varepsilon$, then $a_j \neq \varepsilon$;
- if $u_1 = \varepsilon$, then $a = v_1$ and $b = w_1$;
- if $u_1 \neq \varepsilon$, then there exists a sequence $1 \leq i_1 \leq j_1 < i_2 \leq j_2 < \dots < i_k \leq j_k \leq n$ such that: $u_1 = a_{i_1} \dots a_{j_1} = b_{i_1} \dots b_{j_1}$, $u_2 = a_{i_2} \dots a_{j_2} = b_{i_2} \dots b_{j_2}$, $u_k = a_{i_k} \dots a_{j_k} = b_{i_k} \dots b_{j_k}$;
- if $v_1 = w_1 = \varepsilon$, then $i_1 = 1$;
- if $v_{k+1} = w_{k+1} = \varepsilon$, then $j_k = n$;
- if $k \geq 2$, then for any $i \in \{2, \dots, k\}$ we have $v_i \neq \varepsilon$ or $w_i \neq \varepsilon$.

In this case $l(u_1) + l(u_2) + \dots + l(u_k) = |\{i : a_i = b_i\}|$.

The above decompositions forms are called *parallel decompositions* of strings a and b [55, 56, 57]. For any parallel decompositions $a = v_1u_1 \dots v_ku_kv_{k+1}$ and $b = w_1u_1 \dots w_ku_kw_{k+1}$ the number

$$E(v_1u_1 \dots v_ku_kv_{k+1}, w_1u_1 \dots w_ku_kw_{k+1}) \\ = \sum_{i \leq k+1} \{\max\{l(v_i), l(w_i)\}\} = d_H(x_1x_2 \dots x_n, y_1y_2 \dots y_n)$$

is called the efficiency of the given parallel decompositions. The number $E(a, b)$ is equal to the minimum of efficiency values of all parallel decompositions of the strings a, b and is called the *common* *efficiency of the strings a,b*. It is obvious that $E(a, b)$ is well determined and $E(a, b) = d_G(a, b)$. We say that the parallel decompositions $a = v_1u_1v_2u_2 \ldots v_ku_kv_{k+1}$ and $b = w_1u_1w_2u_2 \ldots w_ku_kw_{k+1}$ are optimal if the following equality holds:

$$E(v_1u_1v_2u_2 \ldots v_ku_kv_{k+1}, w_1u_1w_2u_2 \ldots w_ku_kw_{k+1}) = E(a, b).$$

This type of decompositions are associated with the problem of approximate string matching [35]. If the decompositions $a = v_1u_1 \ldots v_ku_kv_{k+1}$ and $b = w_1u_1 \ldots w_ku_kw_{k+1}$ are optimal and $k \geq 2$, then we may consider that $u_i \neq \varepsilon$ for any $i \leq k$.

Any parallel decompositions $a = a_1a_2 \ldots a_n = v_1u_1v_2u_2 \ldots v_ku_kv_{k+1}$ and $b = b_1b_2 \ldots b_n = w_1u_1w_2u_2 \ldots w_ku_kw_{k+1}$ generate a common sub-sequence $u_1u_2 \ldots u_k$. The number

$$m(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = l(u_1) + l(u_2) + \ldots + l(u_k)$$

is the *measure of similarity* of the decompositions [7, 37]. There exist parallel decompositions $a = v_1u_1v_2u_2 \ldots v_ku_kv_{k+1}$ and $b = w_1u_1w_2u_2 \ldots w_ku_kw_{k+1}$ for which the measure of similarity is maximal. The maximum value of the measure of similarity of all decompositions is denoted by $m^*(a, b)$. The maximum value of the measure of similarity of all optimal decompositions is denoted by $m^\omega(a, b)$. We can note that $m^\omega(a, b) \leq m^*(a, b)$. For any two parallel decompositions $a = a_1a_2 \ldots a_n$ and $b = b_1b_2 \ldots b_n$ as in [56], we define the *penalty factors* as

$$p_r(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = |\{i \leq n : a_i = \varepsilon\}|, p_l(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = |\{j \leq n : b_j = \varepsilon\}|,$$
$$p(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = |\{i \leq n : a_i = \varepsilon\}| + |\{j \leq n : b_j = \varepsilon\}|$$
$$= p_r(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) + p_l(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n)$$

and

$$M_r(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = m(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) - p_r(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n),$$
$$M_l(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = m(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) - p_l(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n),$$
$$M(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = m(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) - p(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n)$$

as the *measures of proper similarity*.

The number $d_H(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = |\{i \leq n : a_i \neq b_i\}|$ is the Hamming distance between decompositions and it is another type of penalty. We have that

$$p(a_1 \ldots a_n, b_1 \ldots b_n) \leq d_H(a_1 \ldots a_n, b_1 \ldots b_n).$$

The assertions from the following theorem establish the main results.

Theorem 3.3.1. *Let a and b be two non-empty strings,* $a = a_1a_2 \ldots a_n$ *and* $b = b_1b_2 \ldots b_n$ *be the initial optimal decompositions, and* $a = a'_1a'_2 \ldots a'_q$ *and* $b = b'_1b'_2 \ldots b'_q$ *be the second decompositions, which are arbitrary. Denote by*

$$
\begin{aligned}
m &= m(a_1a_2\ldots a_n, b_1b_2\ldots b_n), & m' &= m(a'_1a'_2\ldots a'_q, b'_1b'_2\ldots b'_q),\\
p &= p(a_1a_2\ldots a_n, b_1b_2\ldots b_n), & p' &= p(a'_1a'_2\ldots a'_q, b'_1b'_2\ldots b'_q),\\
p_l &= p_l(a_1a_2\ldots a_n, b_1b_2\ldots b_n), & p'_l &= p_l(a'_1a'_2\ldots a'_q, b'_1b'_2\ldots b'_q),\\
p_r &= p_r(a_1a_2\ldots a_n, b_1b_2\ldots b_n), & p'_r &= p_r(a'_1a'_2\ldots a'_q, b'_1b'_2\ldots b'_q),\\
r &= d_H(a_1a_2\ldots a_n, b_1b_2\ldots b_n), & r' &= d_H(a'_1a'_2\ldots a'_q, b'_1b'_2\ldots b'_q),
\end{aligned}
$$

$M = m - p$, $M' = m' - p'$, $M_l = m - p_l$, $M'_l = m' - p'_l$, $M_r = m - p_r$, $M'_r = m' - p'_r$.

The following assertions are true:

1. $p' - p = 2(m' - m) + 2(r' - r)$.
2. *If the second decompositions are non optimal, then $M_l > M'_l$ and $M_r > M'_r$.*
3. *If the second decompositions are optimal, then $M_l = M'_l$ and $M_r = M'_r$ and the measures M_l and M_r are constant on the set of optimal parallel decompositions.*
4. *If $m' \geq m$ and the second decompositions are non optimal, then $p' > p$, $p_l\prime > p_l$, $p'_r > p_r$ and $M > M'$.*
5. *If $m' = m$ and the second decompositions are optimal, then $p' = p$, $p_l\prime = p_l$, $p'_r = p_r$ and $M' = M$.*
6. *If $m' \leq m$ and the second decompositions are non optimal, then $m' - r' < m - r$.*

From Assertions of Theorem 3.3.1 it follows that on the class of all optimal decompositions of given two strings:

- the maximal measure of proper similarity is attained on the optimal parallel decomposition with minimal penalties (minimal measure of similarity);
- the minimal measure of proper similarity is attained on the optimal parallel decomposition with maximal penalties (maximal measure of similarity).

For any two non-empty strings there are parallel decompositions with maximal measure of similarity and optimal decompositions on which the measure of similarity is minimal.

We present below an example that illustrates the relations on proper similarities and penalties implied by Assertion 4 of Theorem 3.3.1.

Example 3.3.1. *Let $a = AAAACCC$, $b = CCCBBBB$ be two strings with a and b being their trivial optimal decompositions, and $a' = AAAACCC\varepsilon\varepsilon\varepsilon\varepsilon$, $b' = \varepsilon\varepsilon\varepsilon\varepsilon CCCBBBB$ as their non-optimal decompositions. Then $m' = 3$, $r' = 8$, $p' = 8$,$m = 0$, $r = 7$, and $p = 0$. In this example we have that $-5 = m' - r' > m - r = -7$ and $-5 = m' - p' = M' < M = m - p = 0$.*

The following example shows that there are some exotic non-optimal parallel decompositions $a = a'_1a'_2\cdots a'_q$ and $b = b'_1b'_2\cdots b'_q$, such that for optimal decompositions $a = a_1a_2\cdots a_n$ and $b = b_1b_2\cdots b_n$ we have $m' < m$, $p' < p$, and $M' > M$.

Example 3.3.2. *Let $ABCDEF$ and $CDEFED$ be trivial non-optimal decompositions of strings a, b and $ABCDEF\varepsilon\varepsilon$, $\varepsilon\varepsilon CDEFED$ be their optimal decompositions. Then $m' = 1$, $r' = 5$, $p' = 0$, $m = 4$, $r = 4$, and $p = 4$. We have that $m' - p' = M' > M = m - p$, and $m' - r' < m - r$.*

The above examples show that Theorem 3.3.1 cannot be improved in the case of $m' < m$.

Decompositions with minimal penalty and maximal proper similarity are of significant interest. Moreover, if we solve the problem of text editing and correction, the optimal decompositions are more favorable. Therefore, the optimal decompositions are the best parallel decompositions and we may solve the string match problems only on class of optimal decompositions.

In virtue of Theorem 3.3.1, various applications of distances on monoids of strings can be used in solving problems from distinct scientific fields. As an example, the study of the measure of proper similarity is approached from a new perspective.

Remark 3.5.1. *Our algorithms are effective for any quasi-metric on $\bar{A}$. Some authors consider the possibility to define the generalized Levenshtein metric with distinct values $\rho(a, b)$ and $\rho(b, a)$. It is necessary to require that $\rho(a, b)$ is a quasi-metric. In other cases we may obtain some confusions as will be seen from the next example.*

Example 3.5.4. *Let $A = \{a, b\}$, $\bar{A} = \{\varepsilon, a, b\}$. The following table defines the distance ρ on $\bar{A}$:*

0	0	1	ε
1	0	0	a
0	1	0	b
ε	a	b	$y \diagdown x$

In this example we have $0 = \rho(a, b) + \rho(b, \varepsilon) < \rho(a, \varepsilon) = 1$ and:

1. *for $u = aba$, $v = ba$ we get $\bar{\rho}(u, v) = \bar{\rho}(v, u) = 0$,*
2. *for $u = a$, $v = b$ we get $\bar{\rho}(u, v) = \bar{\rho}(v, u) = 0$, when $\rho(v, u) = 1$.*

Example 3.5.5. *Let us examine the example from [37] in the context of the results achieved. We have strings $a = AJCJNRCKCRBP$ and $b = ABCNJROCLCRPM$ for which there are eight pairs of optimal decompositions. We present two of them, the shortest and the longest:*

$$\begin{pmatrix} A \\ A \end{pmatrix} \begin{matrix} J \\ B \end{matrix} \begin{pmatrix} C \\ C \end{pmatrix} \begin{matrix} \varepsilon \\ N \end{matrix} \begin{pmatrix} J \\ J \end{pmatrix} \begin{matrix} N & R \\ R & O \end{matrix} \begin{pmatrix} C \\ C \end{pmatrix} \begin{matrix} K \\ L \end{matrix} \begin{pmatrix} C & R \\ C & R \end{pmatrix} \begin{matrix} B & P \\ P & M \end{matrix}$$

$$\begin{pmatrix} A \\ A \end{pmatrix} \begin{matrix} J \\ B \end{matrix} \begin{pmatrix} C \\ C \end{pmatrix} \begin{matrix} J \\ \varepsilon \end{matrix} \begin{pmatrix} N \\ N \end{pmatrix} \begin{matrix} \varepsilon \\ J \end{matrix} \begin{pmatrix} R \\ R \end{pmatrix} \begin{matrix} \varepsilon \\ O \end{matrix} \begin{pmatrix} C \\ C \end{pmatrix} \begin{matrix} K \\ L \end{matrix} \begin{pmatrix} C & R \\ C & R \end{pmatrix} \begin{matrix} B \\ \varepsilon \end{matrix} \begin{pmatrix} P \\ P \end{pmatrix} \begin{matrix} \varepsilon \\ M \end{matrix}$$

For the first pair we have $\rho^* = 7$, $m = 6$, $p = 1$, and $M = 5$. For the second pair we have $\rho^* = 7$, $m = 8$, $p = 5$, and $M = 3$. Our algorithms allow us to calculate all optimal decompositions with distinct measure of similarity. Authors from [37] prefer the second pair of decomposition since it has maximal possible measure of similarity. We consider more preferable the first pair, which has the maximal proper similarity.

In fact, for any two non-empty strings there exist the parallel decompositions with maximal measure of similarity and the optimal decompositions on which measure of similarity is minimal. The following example shows that there exist some exotic non optimal parallel decompositions $a = a'_1 a'_2 \cdots a'_q$ and $b = b'_1 b'_2 \cdots b'_q$, such that for optimal decompositions $a = a_1 a_2 \cdots a_n$ and $b = b_1 b_2 \cdots b_n$ we have $m' < m$, $p' < p$ and $M' > M$.

Example 3.5.6. *Let $a = ABCDEF$ and $b = CDEFED$ be trivial non optimal decompositions of strings a, b, and $a = ABCDEF\varepsilon\varepsilon$ and $b = \varepsilon\varepsilon CDEFED$ be their optimal decompositions. Then $m' = 1$, $r' = 5$, $p' = p'_l = p'_r = 0$ and $m = 4$, $r = 4$, $p = 4$, $p_l = p_r = 2$. In this example we have that $M'_l = M'_r = M' = m' - p' = 1 - 0 = 1 > 0 = 4 - 4 = m - p = M$, $m' - r' = -4 < 0 = m - r$, $M_l = 4 - 2 = 2 > 1 = M'_l$, $M_r = 4 - 2 = 2 > 1 = M'_r$.*

Example 3.5.7. *Let $a = AAAACCC$ and $b = CCCBBBB$ be trivial optimal decompositions of strings a, b, and $a = AAAACCC\varepsilon\varepsilon\varepsilon\varepsilon$ and $b = \varepsilon\varepsilon\varepsilon\varepsilon CCCBBBB$ be their non-optimal decompositions. Then $m' = 3$, $r' = 8$, $p' = 8$ and $m = 0$, $r = 7$, $p = p_l = p_r = 0$. In this example we have that $-5 = m' - r' > m - r = -7$ and $-5 = m' - p' < m - p = 0$.*

The above examples show that Theorem 3.3.1 cannot be improved in the case of $m' < m$. The examined properties and the results obtained in this chapter are published in the articles [46, 47, 51, 55, 56, 57, 62, 63] and serve as a foundation for the next chapter. The mentioned results can also be applied in various problems related to similarity between sequences of characters.

Chapter 4, Geometrical and topological aspects of information analysis, is the final chapter of the thesis, and focuses on the applicative part of the theoretical results obtained in previous chapter. More specifically, the problem of constructing the weighted means and the bisector sets of a pair of strings is solved in this chapter. This result relies strongly on the properties of the parallel optimal decompositions. It is proved that any element of the set of the weighted means is generated by some parallel optimal decompositions. This strong property, as well as its converse, are summarized in the following two theorems.

Theorem 4.1.1. *Any fixed parallel d-optimal decompositions of a pair given strings $a, b \in L(A)$ generate weighted means, simultaneously with their equivalent representations, which form parallel d-optimal decompositions with the fixed representations of the given strings.*

Corolarry 4.1.1. *Any weighted mean of a fixed pair of strings is generated by some of their optimal parallel decompositions.*

As a special case, when d_G (Graev distance) is discrete metric and $d_G(a, b)$ is an even number, the weighted mean equally distant from the strigns a and b is the median of the segment with ends in a and b. The algorithm for constructing the medians of a pair strings is presented below.

The chapter continues with the study of the question of the convexity of the set of the weighted means. The theorems that follow, approach this questions using Hamming and Graev distances.

Algorithm 5: Medians of OPD of x and y:
Given $x, y \in L(\bar{A})$ construct $m \in L(\bar{A})$, s.t. $d^*(x,m) = d^*(m,y)$.

Data: $x = x_1 x_2 \ldots x_n$, $y = y_1 y_2 \ldots y_m$.
Result: Set M of median strings m.
$d := d^*(x, y)$;
if *d is odd* **then**
  return "distance $d^*(x,y)$ is odd, set M is an empty set."
`// Generate Optimal Parallel Decompositions of strings x,y`
$OPD(x, y)$:= BuildOPD(x,y);
$I = \{i : 1 \le i \le l^*(x')\}$;
foreach $(x', y') \in OPD(x, y)$ **do**
  $I_1 = \{i : x'_i = y'_i\}$;
  $I_2 = I \setminus I_1$;
  foreach $I_3 =$ *Choose* $(|I| - d)/2$ *elements from* I_2 **do**
    $m := m_1 m_2 \ldots m_{|I|}$, where $m_i = \begin{cases} x'_i, i \in I_1 \cup I_3 \\ y'_i, otherwise. \end{cases}$
  $M := M \cup \{m\}$;
return M;

Question 1. *Is it true that the set $M_{d_H}(a,b)$ is d_H-convex in $(L^*(A), d_H)$ for any $a, b \in L^*(A)$?*

Question 2. *Is it true that the set $M_{d_G}(a,b)$ is d_G-convex in $(L^*(A), d_G)$ for any $a, b \in L^*(A)$?*

Theorem 4.2.1. *The set $M_{d_H}(a,b)$ is d_H-convex in $(L^*(A), d_H)$ for any $a, b \in L^*(A)$.*

Theorem 4.2.2. *There exists a finite alphabet A and two strings $a, b \in L(A)$ for which the set $M_{d_G}(a,b)$ is not d_G-convex.*

The following example shows such strings $a, b \in L(A)$ for which $M_{d_G}(a,b)$ is not d_G-convex.

Example 4.2.5. *Let $A = \{B, C, D, J, K, L, M, N, O, P, Q, R\}$,*

$$a = DJCJNRCKCRBP,\ b = DBCNJROCLCRPM,$$
$$a' = DJCNJNRCKCRBP,\ b' = DBCJNJROCLCRBPM,$$
$$c = DJCJNJNROCKCRBPM,\ m = QQQQQQQ.$$

Consider the strings amb, bma, $a'ma'$, $b'mb'$ and cmc. We obtain the following:

$$d_G(amb, bma) = 14,\ d_G(amb, a'ma') = d_G(a'ma', bma) = 7,$$
$$d_G(amb, b'mb') = d_G(b'mb', bma) = 7,\ d_G(a'ma', b'mb') = 12,$$
$$d_G(a'ma', cmc) = d_G(cmc, b'mb') = 6,\ d_G(amb, cmc) = d_G(cmc, bma) = 9.$$

Hence $a'ma'$, $b'mb'$ are from the middle of the segment $M_{d_G}(amb, bma)$, the string cmc is from the middle of the segment $M_{d_G}(a'ma', b'mb')$, but $cmc \notin M_{d_G}(amb, bma)$.

The chapter continues with the theorem about the methods of construction of the elements of the bisector set of two strings, i.e. the strings which are equally distant from a and b.

Theorem 4.3.1. *Let $a = a_1a_2...a_n$ and $b = b_1b_2...b_n$ be two strings from $L^*(A)$. There exist methods to construct elements $c = c_1c_2...c_n \in \bar{B}_{d_H}(a, b)$.*

The chapter concludes with the study of the image processing methods using the notions of scattered and digital spaces. One of the results of this study establishes that the Khalimsky topology is the minimal digital topology in the class of all symmetrical topologies on the discrete line $\mathbb{Z}$. We say that the topology $\mathcal{T}$ on $\mathbb{Z}$ is symmetric if $(\mathbb{Z}, \mathcal{T})$ is a scattered Alexandroff space, the set $\{0\}$ is not open in $(\mathbb{Z}, \mathcal{T})$ and for any $n \in \mathbb{Z}$ the mapping $S_n : \mathbb{Z} \to \mathbb{Z}$, where $S_n(x) = 2n - x$ for each $x \in \mathbb{Z}$, is a homeomorphism.

It is well known that distinct algebraic and topological structures have been introduced to accommodate the needs of information theories. In the process of studying of the continuous objects by the computer methods, they are approximated by finite objects or by digital images [1, 9, 11, 56, 30, 38, 39, 41].

Digital image processing is a process which from a topological point of view may be described in the following way:

1. Fix an infinite space X (a continuous image of the original) and a property $\mathcal{P}$ of subspaces of the space X.

2. By some procedure we construct a number $n \in \omega$, a finite subset $H = \{h_i : i \in \omega(n)\} \subset \mathbb{Z}$ of levels and a finite family $\{G_i : i \in \omega(n)\}$ of open non-empty subsets of the space X with the properties:

- $G_i \cap G_k = \emptyset$ for all $0 \leq i < k \leq n$;

- for any $i \in \omega(n)$ and each $x \in G_i$ there exists an open subset $G(x)$ such that $x \in G(x) \subset G_i$ and $G(x)$ is a subset with the property $\mathcal{P}$ in X;

- the set $G = \{G_i : i \in \omega(n)\}$ is dense in X.

The set G is the $\mathcal{P}$-kernel and $X \setminus G$ is the $\mathcal{P}$-residue of the space X.

3. The intensity mapping $I_{\mathcal{P}} : X \to \omega \subseteq H$ is determined with the property: $I_{\mathcal{P}}(x) = maximal\{h_i : x \in cl_X G_i\}$ for each $x \in X$. We have $G_i \subset I_{\mathcal{P}}^{-1}(h_i)$ for each $i \in \omega(n)$.

4. On H is determined a digital topology for which the mapping $I_{\mathcal{P}}$ is continuous.

5. By some procedure we construct a finite T_0-space K and for any $x \in X$ we determine a non-empty subset $D_{\mathcal{P}}(x)$ of K such that:

- for any $c \in K$ the set $X(c) = \{x \in X : c \in D_{\mathcal{P}}(x)\}$ is closed and is called a $\mathcal{P}$-cell of X;

- for any $c \in K$ there exist $i \in \omega(n)$ and an open non-empty subset $X'(c) \subset G_i$ such that $X(c) = cl_X X'(c)$.

The results of the research presented in this chapter, along with the results discussed in the previous chapters fully cover the research goals stated in the first chapter. The results from this chapter are published in the articles [49, 50, 60, 61] and can be applied in the study of various theoretical as well as practical problems.

4. GENERAL CONCLUSIONS

The research carried out within the Ph.D. thesis "Distances on Free Monoids and Their Applications in Theory of Information" fully corresponds to the goals and the objectives set out in the introduction chapter.

The study of the results obtained permit to highlight the following general results:

1. It was established that for any non-Burnside quasivariety $\mathcal{V}$ and any quasi-metric ρ on a set X with basepoint p_X on free monoid $F^a(X, \mathcal{V})$ there exists a unique stable quasi-metric $\hat{\rho}$ with the properties:
   (a) $\rho(x, y) = \hat{\rho}(x, y)$ for all $x, y \in X$;
   (b) If d is an invariant quasi-metric on $F^a(X, \mathcal{V})$ and $d(x, y) \leq \rho(x, y)$ for all $x, y \in X$, then $d(x, y) \leq \hat{\rho}(x, y)$ for all $x, y \in F^a(X, \mathcal{V})$;
   (c) If ρ is a metric, then $\hat{\rho}$ is a metric as well;
   (d) If $Y \subseteq X$, $d = \rho|Y$ and $\hat{d}$ is the maximal invariant extension of d on $F^a(Y, \mathcal{V})$, then $F^a(Y, \mathcal{V}) \subseteq F^a(X, \mathcal{V})$ and $\hat{d} = \hat{\rho}|F^a(Y, \mathcal{V})$;
   (e) For any quasi-metric ρ on X and any points $a, b \in F^a(X, \mathcal{V})$ there exists $n \in N$ and representations $a = a_1a_2...a_n$, $b = b_1b_2...b_n$, such that $a_1, b_1, a_2, b_2, ..., a_n, b_n \in X$ and $\hat{\rho}(a, b) = \sum\{\rho(a_i, b_i) : i \leq n\}$. [62]

2. The method of extension of quasi-metrics on free monoids in the complete non-Burnside quasivariety of topological monoids permit: to construct distinct admissible topologies of $F^a(X, \mathcal{V})$ for any T_0-space X, to prove that the free topological monoid $F^a(X, \mathcal{V})$ exists for any space X, to establish that the free topological monoid $F(X, \mathcal{V})$ is abstract free, i.e. is canonically isomorphic with the abstract free monoid $F^a(X, \mathcal{V})$ [54, 57, 62].

   This fact solves problems posed by A. I. Maltsev for free universal topological algebras [33]. Similar results were obtained for quasivarieties of semi-topological monoids as well [62].

3. It was proved that if $\mathcal{V}$ is a complete non-Burnside quasivariety of topological monoids, then X is an Alexandroff space if and only if $F(X, \mathcal{V})$ is an Alexandroff space, and X is a digital space if and only if $F(X, \mathcal{V})$ is a digital space [61].

   We mention that conclusions 1, 2 and 3 do not hold for complete Burnside quasivarieties.

4. Based on distance extension methods, the notions of parallel decompositions and the measure of similarity were introduced in the space of strings [63]. Theorem describes the relationships between measure of similarity, penalty and optimality of parallel decompositions [56, 58, 59].

5. Different interesting relations between Hamming, Levenshtein and Graev distances were established on $L(A)$ [46, 47, 48, 55].

6. It was proved that on the class of all optimal decompositions of given two strings the maximal measure of proper similarity is attained on the optimal parallel decomposition with minimal

penalties (minimal measure of similarity), and the minimal measure of proper similarity is attained on the optimal parallel decomposition with maximal penalties (maximal measure of similarity) [51, 63].

7. Algorithms were proposed for constructing the elements of the sets of weighted means $M_{d_G}(a, b)$ and bisector $B_{d_G}(a, b)$ of a given pair of strings a and b [49, 64]. It was illustrated how to use optimal parallel decompositions to generate elements of $M_{d_G}(a, b)$, $B_{d_G}(a, b)$, and the set of midpoints between a and b [50, 53, 64].
8. It was proved that any weighted mean of a pair of strings is generated by some of their optimal parallel decompositions [64]. It was also proved that the set $M_{d_G}(a, b)$ is not convex [52].
9. Algorithms for digital image processing were elaborated using the properties of scattered and digital topologies, and it was established that the Khalimsky topology is the minimal digital topology in the class of all symmetrical topologies on the discrete line $\mathbb{Z}$ [60, 61, 65].

Advantages and value of thesis results. The proposed elaborations have a significant scientific value due to their high degree of novelty and originality. The scientific results in this thesis have a theoretical and applicative value in domains of algebra, topology and theoretical computer science. For example, the methods of extensions of pseudo-quasimetrics that can be used for construction of special topologies on free monoids. The methods of parallel decompositions, measure of similarity, efficiency and penalty can be applied in text analysis problems.

Recommendations. The results obtained can be used in various fields and may have practical applications in algebra and theory of information. Based on the above conclusions, we recommended the following:

- there is a special interest in investigating quasimetrics on the space of free monoids, as extensions of quasimetrics with particular properties on an alphabet. For instance, as it was proved, quasimetrics are strictly invariant on rigid quasivarieties. This is usual for groups, but it is very rare for semigroups and monoids;
- the results research can be continued both from algebraic and applicative points of view. Researching metrics on monoids is of particular interest;
- the results obtained with optimal parallel decompositions can be used in the domain of sequences alignment;
- the new algorithm proposed for weighted means construction can be more effective because it takes into consideration the empty symbol, and generates more elements of the M_{d_G} set than the classical algorithms. This fact, in its turn, can be useful in the context of information communication through the channel with noise, or text editing/correction software, where the loss of information takes place;
- algorithms for generating weighted means and bisectors of strings can be applied in the domain of data analysis and clustering algorithms. For instance, the geometrical centroid of

a set of elements can be calculated as the intersection of the bisectors of elements;

- further research can be continued with the study of algorithms and properties of optimal parallel decompositions of three and more strings;
- thesis contents can serve as a platform for university facultative and optional courses.

References

[1] Abramsky S., Jung A. *Domain Theory in Handbook of Logic in Computer Science*, vol. 3, edited by S. Abramsky, Dov. Gabbay, and T.S.E. Maibaum eds., Oxford University Press, 1994, p. 1–168. ISBN 0-19-853762-X.

[2] Alexandroff P. S. *Diskrete Raume*. Mat. Sb. (N.S.) (in German). vol. 2, 1937, p. 501–518.

[3] Arbib M. A. *The Algebraic Theory of Machines, Languages, și Semigroups*, Academic Press, New York, 1968.

[4] Arhangelskii A. V. *Mappings și spaces*. Uspehi Matem. Nauk 21 (1966), vyp. 4, p. 133–184 (English translation: Russian Math. Surveys, vol. 21, 1966, no. 4, p. 115–162).

[5] Arhangelskii A.V., Choban M.M., Kenderov P.S. *Topological games and topologies on groups*, Math. Maced., nr. 8, 2010, p. 1–19. ISSN 1409 9721.

[6] Arnautov V. I., Ermakova G. N. *Lattice of all topologies of countable module over countable rings*, Buletinul Academiei de Ştiinţe a Republicii Moldova, Matematica, nr.2 (81), 2016, p. 63–70. ISSN 1024-7696.

[7] Barakhnin V., Nekhaeva V., Fedotov A. *Similarity Determination for Clustering Textual Documents*. In: Wolff K.E., Palchunov D.E., Zagoruiko N.G., Andelfinger U. (eds) Knowledge Processing and Data Analysis. KPP 2007, KONT 2007. Lecture Notes in Computer Science, vol 6581. Springer, Berlin, Heidelberg. ISBN 978-3-642-22139-2.

[8] Belousov V.D. *Foundations of the theory of quasigroups și loops* (in Russian), Izdatyelstvo, Moscow, 1967.

[9] Bentley J.L., Weide B.W., Yao A.C. *Optimal expected-time algorithms for closest point problems*, ACM Trans. Math. Software vol. 6, 1980, p. 563–580. ISSN 0098-3500.

[10] Birkhoff G. *A note on topological groups*, Compositio Math., vol. 3, 1936, p. 427–430.

[11] Blanck J. *Domain representations of topological spaces*, Theoretical Computer Science, vol. 247, 2000 , p. 229–255. ISSN 0304-3975.

[12] Choban M. *General conditions of the existence of free algebras*, Acta Comment. Univ. Tartuensis, vol. 836, 1989, p. 157–171.

[13] Choban M. *Some topics in topological algebra*, Topol. Appl., vol. 54, 1993, nr. 54, p. 183–202.

[14] Choban M. *On the theory of topological algebraic systems*, Trans. Moscow Math. Soc., vol. 48, 1986, p. 115–159 (Russian original: Trudy Mosk. Matem. Ob-va, vol. 48, 1985, p. 106–149).

[15] Choban M. *The theory of stable metrics*, Math. Balkanica, vol. 2, 1988, p. 357–373. ISSN 0205-3217.

[16] Choban M. M., Chiriac L. L. *On free groups in classes of groups with topologies*, Bul. Acad. Ştiinţe Repub. Moldova, Matematica, 2013, nr 2-3, p. 61–79. ISSN: 0236-3089.

[17] Clifford A.H., Preston G.B. *The algebraic theory of semigroups*, Math Surveys No. 7, Amer. Math. Soc. Providence, vol. 1, 1961, vol. 2, 1967. ISSN 1088-6826.

[18] Cojocaru S., Petic M., Titchiev I., *Adapting Tools for Text Monitoring și for Scenario Analysis Related to the Field of Social Disasters*, International Journal of Computer, Electrical, Automation, Control și Information Engineering, vol. 10, nr. 10, 2016, p. 1601–1607. ISSN 1740-7524.

[19] Dumitrascu S. S., Choban M. *On free topological algebras with a continuous signature*. In: Algebraic și topological systems. Matem. Issledovania, vol. 65, 1982, p. 27–54.

[20] Eilenberg S., Wright J. B. *Automata in General Algebras*, Information și Control, vol. 11, 1967, p. 452–470.

[21] Ershov Yu. L. *Definability și Computability*, Springer, 1996. ISBN 978-0-306-11039-9.

[22] Frechet M. M. *Sur quelques points du calcul fonctionnel*, Rendic. Circ. Mat. Palermo, vol. 22, 1906, p. 1–74.

[23] Gaindric C., Cojocaru S., Popcova O., Puiu S., Secrieru Iu. *Emergency-SonaRes: A System for Ultrasound Diagnostics Support in Extreme Cases.* In: H.-N.L. Teodorescu et al. (eds.), Improving Disaster Resilience și Mitigation - IT Means și Tools, NATO Science for Peace și Security Series C: Environmental Security, Part I, Fundamentals și Modeling, Chapter 18, 2014, p. 283–292. ISBN 978-94-017-9136-6.

[24] Glushkov V. M. *Abstract Theory of Automata*, Uspehi Matematiceskih Nauk 16:5 (1961), 3-62, (English translation:Russian Mathematical Surveys, 1961, 16:5, 1-53; Romanian translation: Analele Romano-Sovetice, Seria Matematica-Fizica, vol. 2, nr. 41, 1962, p. 3-61).

[25] Graev M. I. *Theory of topological groups. I. Norms și metrics on groups. Complete groups. Free topological groups*. (Russian) Uspehi Matem. Nauk (N.S.) 5, (1950). no. 2(36), 3-56.

[26] Hamming R. W. *Error Detecting și Error Correcting Codes*, Bell System Technical Journal, vol. 29, nr. 2, 1952, p. 147–160. ISSN 0005-8580.

[27] Heath R. W. *On spaces with point countable bases*, Bull. Acad. Polon Sci., Ser Math., vol. 13, nr. 6 , 1965, p. 393–395.

[28] Higgins P. *Categories și groupoids*, Van Nostrand-Reinhold, New York, 1971.

[29] Hofmann K.H., Lawson J.D., Pym J.S. *The analytical și topological theory of semigroups*, de Gruyter, 1990. 400 p. ISBN 978-3-11-085604-0.

[30] Kronheimer E.H. *The topology of digital images*, Topology și its Applications vol. 46, 1992, p. 279–303. ISSN 0166-8641.

[31] Levenshtein V.I. *Binary codes capable of correcting deletions, insertions, și reversals*, DAN SSSR, vol. 163, nr. 4, p. 845–848, 1965 (in Russian) (English translation: *Soviet Physics – Doklady*, vol. 10, nr. 8, p. 707–710, 1966).

[32] Lind D., Marcus B. *An Introduction to Symbolic Dynamics și Coding*, Cambridge University Press, Cambridge, 1995. ISBN 9780511626302.

[33] Maltsev A. I. *Free topological algebras*, Trans. Moscow Math. Soc., vol. 2, nr. 17, 1961, p. 173–200 (Russian original: Izvestia Akad. Nauk SSSR, vol. 21, 1957, p. 171–198).

[34] Maltsev A. I. *On the general theory of algebraic systems*, Mat. Sb. vol. 35, 1954, p. 3–20; English translation: Trans. Amer. Math. Soc., vol. 27, 1963, p. 125–148.

[35] Navarro G. *A guided tour to approximate string matching*, ACM Computing Surveys, vol. 33, no. 1, 2001. p. 31–88. ISSN 0360-0300.

[36] Nedev S.I. *o-metrizable spaces*, Trudy Moskov. Mat.Ob-va, vol. 24 1971, p. 201–236 (English translation: Trans. Moscow Math. Soc., vol 24, 1974, p. 213–247).

[37] Needleman S. B., Wunsch C. D. *A general method applicable to the search for similarities in the amino acid sequence of two proteins*. Journal of Molecular Biology, vol. 48, nr. 3, 1970, p. 443–453. ISSN 0022-2836.

[38] Roy N., Das A. *Topological Structure in Digital Image Processing : A Survey*, International Journal of Scientific Research in Computer Science, Engineering și Information Technology, vol. 2, 2017, nr. 6, p. 481–485. ISSN 2456-3307.

[39] Sandhu R.S. *Lattice-based access control models*, Computer, vol. 26, nr. 11, 1993, p. 9–19. ISSN 0018-9162.

[40] Shcherbacov V. A. *Quasigroup based crypto-algorithms*, 2012, arXiv:1201.3016.

[41] Spreen D. *On Some Decision Problems in Programming*, Information și Computation, vol. 122, 1995, nr. 1, p. 120–139. ISSN 0890-5401.

[42] Swierczkowski S. *Topologies in free algebras*, Proceed. London Math. Soc., vol. 15, nr. 55, 1964, p. 566–576.

[43] Syrbu P., Izbash V. *Recursively differentiable quasigroups and complete recursive codes*, Commentat. Math. Univ. Carol. 45, nr. 2, 2004, p. 257–263.

[44] Wallace A. D. *The structure of topological semigroups*, Bull. Amer. Math. Soc. vol. 61, 1955, nr. 2, p. 95–112. ISSN 1088-9485.

[45] Whitehead A. N. *A Treatise on Universal Algebra*, 1898. ISBN 978-1429700320.

AUTHOR PUBLICATIONS ON THESIS TOPIC

[46] Budanaev I. *On Hamming Type Distance Functions*, CAIM 2016: The 24^{th} Conference on Applied și Industrial Mathematics, Craiova, Romania, September 15-18, 2016, Book of Abstracts, Editura SITECH, Craiova, p. 15–16. ISSN 2537 - 2688.

[47] Budanaev I. *On Hamming Type Distance Functions*, Romai Journal 2016, vol. 12, nr. 2, p. 25–32. ISSN 1841-5512.

[48] Budanaev I. *About Distance Functions of Hamming Type*. In: Tendinţe Contemporane ale Dezvoltării Ştiinţei: Viziuni ale Tinerilor Cercetători. Materialele Conferinţei Ştiinţifice a Doctoranzilor, Universitatea AŞM, Ediţia a V-a. Chisinău, 2016, p. 297–301.

[49] Budanaev I. *About the Construction of the Bisector of Two Strings*, Romai Journal, vol.13 n.2, 2017, p. 1 – 11. ISSN 1841-5512.

[50] Budanaev I. *On the Bisector of a Pair of Strings*, Conference of Mathematical Society of Moldova: Proceedings CMSM4 2017, June 25 – July 2, 2017, Chisinau, Republic of Moldova, p. 475–478, 2017, ISBN: 978-9975-71-915-5.

[51] Budanaev I. *Parallel Decompositions și The Weighted Mean of a Pair of Strings*, International Conference "Contemporary Trends in Science Development: Visions of Young Researchers", 6th Edition, Chisinau, 15 June 2017, Proceedings p. 7–11.

[52] Budanaev I. *About Non-Convexity of the Weighted Mean of a Pair of Strings*, International Conference “Contemporary Trends in Science Development: Visions of Young Researchers”, 7th Edition, Chisinau, 15 June 2018, Proceedings vol. 1, p. 6–10.

[53] Budanaev I. *On the Midset of Pairs of Strings*, International Conference on Mathematics, Informatics și Information Technologies Dedicated to the Illustrious Scientist Valentin Belousov, MITI 2018, Communications, p. 134–135. ISBN 978-9975-3214-7-1.

[54] Budanaev I., Choban M. *Invariant Distances on Free Semigroups și Their Applications*. Rezumatele Comunicărilor, A XX-a Conferinţă Anuală a Societăţii de Ştiinţe Matematice din România, Baia Mare, 19-22 mai 2016, p. 17–18.

[55] Choban M., Budanaev I. *Distances on Monoids of Strings și Their Applications*, Proceedings of the Conference on Mathematical Foundations of Informatics MFOI2016, July 25-29, 2016, Chisinau, Republic of Moldova, 2016, p. 144–159. ISBN 978-9975-4237-4-8.

[56] Choban M., Budanaev I. *About Applications of Distances on Monoids of Strings*, Computer Science Journal of Moldova, 2016, nr. 3, p. 335–356. ISSN 1561-4042.

[57] Choban M., Budanaev I. *On the Theory of Free Topological Monoids și its Applications*, International Conference "Mathematics și Information Technologies: Research și Education" (MITRE-2016). Abstracts. Chisinau, June 23-26, 2016, p. 21–22. ISBN 978-9975-71-794-6.

[58] Choban M., Budanaev I. *Parallel Decompositions of Pairs of Strings și Their Applications*, Presented in plenary at CAIM 2017: The 25^{th} Conference on Applied și Industrial Mathematics, Iaşi, Romania, September 14-17, 2017, Book of Abstracts, p. 64–65.

[59] Choban M., Budanaev I. *Measures of Similarity on Monoids of Strings*, Conference on Mathematical Foundations of Informatics: Proceedings MFOI 2017, November 9-11, 2017, Chisinau, Moldova, Institute of Mathematics și Computer Science, p. 51–58. ISBN 978-9975-4237-6-2

[60] Choban M., Budanaev I. *About applications of topological structures in computer sciences și communications*, Acta Et Commentationes, Ştiințe Exacte și ale Naturii, Revistă Ştiințifică, nr.2 (4), 2017, p. 45–59. ISSN 2537-6284.

[61] Choban M., Budanaev I. *Scattered și Digital Topologies in Image Processing*, Proceedings of the Conference on Mathematical Foundations of Informatics, MFOI 2018, July 2-6, 2018, Chisinau, Republic of Moldova, Chisinau, p. 21–40, 2018. ISBN: 978–9975–4237–7–9.

[62] Choban M., Budanaev I. *Distances on Free Semigroups și Their Applications*, Buletinul Academiei de Ştiinţe a Republicii Moldova. Matematica, nr. 1, vol. 86, 2018, p. 92 – 119. ISSN 0236-3089.

[63] Choban M., Budanaev I. *Efficiency și Penalty Factors on Monoids of Strings*, Computer Science Journal of Moldova, vol.26, nr.2 (77), 2018, p. 99 – 114. ISSN 1561-4042.

[64] Choban M., Budanaev I. *About the Construction of the Weighted Means of a Pair of Strings*, Romai Journal, vol.14, nr.1, 2018, p. 73 – 87. ISSN 1841-5512.

[65] Choban M., Budanaev I. *Scattered și Digital Topologies in Information Sciences*, The 26th Conference on Applied și Industrial Mathematics, CAIM 2018, Septmeber 20–23, 2018, Chisinau, Book of Abstracts, p. 9 – 11.

ANNOTATION

of the thesis **"Distances on Free Monoids și Their Applications in Theory of Information"**, submitted by Budanaev Ivan for Ph.D. degree in Mathematics, specialty 111.03 - Mathematical Logic, Algebra și Number Theory.

The thesis was elaborated in Moldova State University "Dimitrie Cantemir", Chişinău, 2019.

Thesis structure: the thesis is written in English și consists of: introduction, four chapters, general conclusions și recommendations, 200 bibliography titles, 116 pages of main text. The obtained results were published in 20 scientific papers.

Keywords: Alexandroff space, quasivariety of topological monoids, free monoids, invariant distance, quasi-metric, Levenshtein distance, Hamming distance, Graev distance, parallel decomposition, proper similarity, weighted mean, bisector of two strings, convexity, algorithm.

Domain of research: Distances on abstract algebraic structures.

Goals și objectives: The goal of the research is to study the problem of distances on free monoids. To achieve this goal, the following objectives were defined: elaboration of an effective method for extending the quasi-metric on free monoids; development of efficient representations of information for data analysis; implementation of innovative algorithms for solving text sequences problems; describe digital topologies on the discrete line.

The scientific novelty și originality consist in obtaining new theoretical results with applications in computer science. An effective method of distance extension on free monoids was developed, which helped to introduce the concept of parallel representation of information. This has allowed the development of the concepts of efficiency și similarity of the representation of information sequences, as well as the construction of the sets of weighted mean și bisector of strings.

The important scientific problem solved in the research is the development of methods for constructing and studying distances on free monoids, which contribute to obtaining effective methods of representing information, applicable to solving different distance problems.

The theoretical significance is determined by the obtaining of the new results regarding the establishment of the conditions of existence of the extension of the distance on free monoids. The elaborated methods have allowed to approach the problems related to information sequences from a new point of view. New algorithms of constructing strings weighted mean și bisector were proposed. It has been established that the informational segment is not convex.

The applicative value of the paper consists in the use of the obtained theoretical results in the study of symmetric topologies on the digital line, imaging processing și construction of the centroid of a set of strings.

The implementation of the scientific results. The obtained results can be used in scientific research related to data analysis, the study of the efficiency of information representation, digital image processing. They can also be used in development of an optional course for university students related to the study of distances on abstract algebraic structures.

ADNOTARE

la teza **"Distanţe pe Monoizi Liberi şi Aplicaţiile lor în Teoria Informaţiei"**, înaintată de către Budanaev Ivan pentru obţinerea titlului de doctor în ştiinţe matematice la specialitatea 111.03 - Logică Matematică, Algebră şi Teoria Numerelor.

Teza a fost elaborată la Universitatea de Stat "Dimitrie Cantemir", Chişinău, anul 2019.

Structura tezei: teza este scrisă în limba engleză şi conţine introducere, patru capitole, concluzii generale şi recomandări, 200 titluri bibliografice, 116 pagini de text de bază. Rezultatele obţinute sunt publicate în 20 lucrări ştiinţifice.

Cuvinte cheie: Spaţiul Alexandrov, cvasivarietate de monoizi topologici, monoizi liberi, distanţă invariantă, cvasimetrică, distanţa Levenshtein, distanţa Hamming, distanţa Graev, descompunere paralelă, similaritate proprie, medie ponderată, bisectoare a două stringuri, convexitate, algoritm.

Domeniul de studiu al tezei: Distanţe pe structuri algebrice abstracte.

Scopul şi obiectivele lucrării. Scopul cercetării constă în studiul problemei extinderei distanţelor pe monoizi liberi. Pentru atingerea acestui scop au fost definite următoarele obiective: elaborarea unei metode eficiente de extindere a cvasimetricei pe monoizi liberi; dezvoltarea reprezentărilor eficiente a informaţiei pentru analiza datelor; implementarea algoritmilor inovativi pentru rezolvarea problemelor secvenţelor de text; descrierea topologiei digitale pe dreapta discretă.

Noutatea şi originalitatea ştiinţifică constau în obţinerea rezultatelor noi de ordin teoretic cu aplicaţii în informatică. A fost elaborată o metodă efectivă de extindere a distanţelor pe monoizi liberi, graţie căreia a fost introdus conceptul de descompunere paralelă a informaţiei. Această a permis dezvoltarea conceptelor de eficienţă şi similaritate a reprezentărei secvenţelor informaţionale, la fel şi construcţia mulţimelor de medii ponderate şi bisectoare a stringurilor.

Problema ştiinţifică importantă soluţioantă constă în elaborarea metodelor de construire și studiere a distanţelor pe monoizi liberi, care contribuie la obţinerea metodelor efective de reprezentare a informaţiei, aplicabile la soluţionarea diferitor probleme referitor la distanţe.

Semnificaţia teoretică este determinată de obţinerea rezultatelor noi ce ţin de stabilirea condiţiilor de existenţă a extinderii distanţei pe monoizi liberi. Metodele elaborate au permis abordarea problemelor legate de secvenţe de informaţie dintr-un nou punct de vedere. Au fost propuşi algoritmi de construcţie a mediilor ponderate şi bisectoarei a perechilor de stringuri. S-a stabilit că segmentul informaţional nu este convex.

Valoarea aplicativă a tezei constă in utilizarea rezultatelor teoretice obţinute la studiul topologiilor simetrice pe dreapta digitală, procesarea imaginelor şi construcţia centrului de greutate a mulţimei de stringuri.

Implementarea rezultatelor ştiinţifice. Rezultatele obţinute pot fi utilizate in cercetări ştiinţifice ce ţin de analiza datelor, studierea eficienţei reprezentării a informaţiei, procesarea digitală a imaginelor. De asemenea, ele pot servi drept suport pentru cursuri universitare opţionale.

АННОТАЦИЯ

диссертации **"Расстояния на свободных моноидах и их приложения в теории информации"**, представленной Иваном Буданаевым на соискание учёной степени доктора математических наук по специальности 111.03 – Математическая Логика, Алгебра и Теория Чисел. Диссертация выполнена в Государственном Университете "Димитрие Кантемир", Кишинёв, 2019 год.

Структура работы: Диссертация написана на английском языке и содержит введение, четыре главы, заключение с рекомендациями, 200 библиографических названия, 116 страниц оцновного текста. Полученные результаты были опубликованы в 20 научных работах.

Ключевые слова: Пространство Александрова, свободные моноиды, инвариантное расстояние, квазиметрика, расстояния Левенштейна, Хэмминга и Граева, параллельное разложение, надлежащее сходство, взвешенное среднее, биссектриса двух строк, выпуклость, алгоритм.

Область исследования: Расстояния на абстрактных алгебраических структурах.

Цель исследования является изучение проблемы расстояний на свободных моноидах, для достижение которого определены следующие задачи: разработка эффективного метода продолжения квазиметрики на свободные моноиды; разработка эффективных представлений информации для анализа данных; внедрение инновационных алгоритмов для решения задач текстовых последовательностей; описание цифровых топологии на дискретной прямой.

Научная новизна и оригинальность заключаются в получении новых теоретических результатов с приложениями в информатике. Разработан эффективный метод продолжения расстояний на свободных моноидах, который позволил ввести концепцию параллельного представления информации, эффективности и сопоставимости информационных последовательностей, а также построить множества взвешенного среднего и биссектрисы строк.

Важной научной задачей, решаемой в исследовании, является разработкаметодов построения и исследования расстояний на свободных моноидах, которые способствуют получению эффективных методов представления информации, применимых для решения задач с расстояниями.

Теоретическая значимость определяется получением новых результатов, касающихся установления условий существования продолжения расстояний на свободных моноидах. Разработанные методы позволили подойти к проблемам, связанным с информационными последовательностями, с новой точки зрения. Предложены новые алгоритмы построения взвешенного среднего и биссектрисы строк. Установлено, что информационный сегмент не является выпуклым.

Прикладная ценность работы заключается в использовании полученных теоретических результатов при исследовании симметричных топологий на цифровой прямой, обработке изображений и построении центроида множества строк.

Реализация научных результатов. Полученные результаты могут быть использованы в научных исследованиях, связанных с анализом данных, изучением эффективности представления информации, цифровой обработкой изображений. Они также могут быть использованы при разработке факультативного курса для студентов университетов, связанного с изучением расстояний на абстрактных алгебраических структурах.

DISTANŢE PE MONOIZI LIBERI
ŞI APLICAŢIILE LOR ÎN TEORIA INFORMAŢIEI

111.03 LOGICA MATEMATICĂ, ALGEBRA ŞI TEORIA NUMERELOR

Rezumatul tezei de doctor în ştiinţe matematice

Aprobat spre tipar:	Formatul hârtiei 60x84 1/16
Hârtie ofset. Tipar ofset.	Tiraj 25 ex.
Coli de tipar: 2	Comanda nr. 1

Chișinău, str. Bănulescu Bodoni, 59.