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

Teza a fost elaborată în Şcoala Doctorală Matematică și Ştiința Informației, Universitatea "Dimitrie Cantemir", în consorțium cu Institulul de Matematică și Informatică "V. A. Andrunachievici", Universitatea de Stat "Alecu Russo" din Bălți, Universitatea de Stat din Tiraspol și Universitatea de Stat "Bogdan Petriceicu Hasdeu" din Cahul.

Conducător de Doctorat:

CIOBAN Mitrofan, academician, doctor habilitat în științe matematice, profesor universitar.

Componenţa Comisiei de Doctorat:

COJOCARU Svetlana, membru corespondent al Academiei de Ştiinţe a Moldovei, doctor habilitat în științe matematice - **președinte**.

ARNAUTOV Vladimir, academician, doctor habilitat în științe matematice, profesor universitar - **referent**.

KENDEROV Petar, academician al Academiei de Ştiinţe a Bulgariei, doctor habilitat în științe matematice, profesor universitar - **referent**.

SÎRBU Parascovia, doctor în științe matematice, conferențiar universitar - **referent**.

CIOBAN Mitrofan, academician, doctor habilitat în științe matematice, profesor universitar - **membru**.

Susţinerea va avea loc la 14.05.2019, ora 14:00, în ședința Comisiei de Doctorat, sala polivalentă a Universităţii de Stat "Dimitrie Cantemir", str. Hânceşti 55/4, Chișinău.

Teza de doctor şi rezumatul pot fi consultate la Biblioteca Ştiinţifică Centrală "Andrei Lupan" (Institut) și pe pagina web a ANACEC (http://www.anacip.md).

Rezumatul a fost expediat la 12.04.2019.

Secretar, Naval Elvira
doctor, conferenţiar cercetător

Conducător ştiinţific, Cioban Mitrofan
academician, profesor universitar

Autor Budanaev Ivan

CUPRINS

CUVINTE CHEIE

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.

1. SCOPUL ȘI OBIECTIVELE CERCETĂRII

Actualitatea și importanța temei de cercetare. Noțiunea de algebră universală a fost introdusă în cartea lui Alfred North Whitehead "A Treatise on Universal Algebra", publicată în 1898 [45]. Între anii 1935 și 1950, Birkhoff a introdus noțiunile de varietăți și cvasivarietăți de algebre universale, algebre libere, congruență de algebre universale, latice de subalgebre, homomorfism. Din cauza celui de-al doilea război mondial, rezultatele publicate de Anatol Maltsev în anii 1938-1946 nu au fost cunoscute până la începutul anilor 50 ai secolului trecut. Raportul plenar a lui Alfred Tarski din 1950 la Congresul Internațional al Matematicienilor la Cambridge a inaugurat o nouă eră. După 1950, au fost studiate diferite aspecte ale teoriei modelelor, cu aplicații neobișnuite în logica matematică, teoria limbajelor, teoria automatelor, cu contribuția următorilor matematicieni: 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. Gluc, 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 și alții (vezi [3, 8, 6, 10, 17, 20, 21, 24, 28, 29, 32, 34, 44]). Diferite aplicații ale algebrei în analiza informației și prelucrarea imaginilor au fost explorate de S. Cojocaru, C. Gaindric, V. Shcherbacov, P. Syrbu, V. Izbash și alții [18, 23, 40, 43]. La rezolvarea multor probleme legate de analiza informației (procesare, digitizare, comparare, clasificare), a devenit necesar să se studieze metrici invariante și topologii pe algebre universale libere. Studiul algebrelor topologice a fost inițiat cu grupuri Lie, grupurilor topologice și spațiilor liniare topologice.

Diferitele tipuri de distanțe au fost examinate de M. Frechet, V. Niemytzcki, P.S. Alexandroff, A.V. Arhangelskii, M.M. Choban, P.Kenderov, R.W. Heath, S. Nedev, W.A. Wilson (a se vedea [22, 2, 5, 27, 4, 15, 36]). În clasa distanțelor, cvasimetricele se deosebesc prin faptul că ele nu sunt simetrice, dar satisfac axioma inegalității triunghiului. Cvasimetricele discrete ne conduc la conceptul de spațiu digital și, mai general, la spațiul Alexandroff. Cercetările lui Hamming [26], Graev [25] și Levenshtein [31] ne conduc la necesitatea de a dezvolta metode de extindere pe monoidul liber $L(A)$ a distanțelor definite pe alfabetul A. Menționăm că este important ca extinderea să fie invariantă. Această problemă este importantă și rămânea nesoluționată, până la momentul dat, pentru orice cvasivarietate a monoizilor topologici.

Faptele menționate determină actualitatea și importanța temei de cercetare.

Teza prezintă rezultatele teoretice ale studiului distanțelor pe structuri algebrice abstracte. Partea aplicativă a cercetării poate fi utilizată în teoria informației, unde este necesar să se definească măsura similarității dintre secvențe de informație și eficiența reprezentării informației. Aceste noțiuni, la rândul lor, pot fi obținute prin aplicarea distanței dintre secvențele de informație.

Scopul și obiectivele cercetării. Scopul cercetării constă în studiul problemei extinderii distanţelor pe monoizi liberi. Printre obiectivele principale ale lucrării se număra studiul proprietăților acestor distanțe pe monoidul liber $L(A)$, de a determina dacă există o relație între ele și de a găsi aplicațiile rezultatelor obținute. Pentru atingerea acestui scop au fost formulate următoarele obiective:

- determinarea condițiilor de extindere a cvasimetricei ρ pe A și orice cvasivarietate $\mathcal{V}$ de monoizi topologici la o cvasimetrică invariantă ρ^* pe monoidul liber $F^a(A, \mathcal{V})$;
- elaborarea unei metode eficiente de extindere a cvasimetricei ρ pe monoizi liberi;
- determinarea condițiilor de existență a topologiilor invariante pe monoidul liber $F^a(A, \mathcal{V})$ care sunt extinderi ale topologiei date pe A;
- stabilirea relațiilor între distanțele Hamming, Levenshtein și Graev;
- dezvoltarea reprezentărilor și descompunerilor eficiente ale perechilor de șiruri de caractere;
- implementarea algoritmilor inovativi pentru rezolvarea problemelor secvențelor de text;
- descrierea prelucrării imaginilor din punct de vedere topologic;
- analiza proprietăților topologiilor digitale pe linia discretă.

Obiectivele stabilite conduc la rezultate teoretice importante. Astfel, a fost introdusă noțiunea de descompunere paralelă a unei perechi de șiruri de caractere. Pe baza acestui fapt, obiectivele de cercetare extinse au vizat studiul măsurii similarității proprii a unei perechi de șiruri de caractere, algoritmii de construcție a mulțimelor de medii ponderate și bisectoarei unei perechi de stringuri, precum și examinarea și soluționarea problemei de convexitate a mulțimilor menționate.

2. METODOLOGIA CERCETĂRII ȘTIINȚIFICE

Studiul prezent se desfășoară în domeniul teoriilor algebrice și topologice, iar metodologia utilizată se bazează pe aplicarea metodelor teoriei monoizilor, a spațiilor cu distanțe, a lingvisticii teoretice, a teoriei algoritmilor și a teoriei sistemelor informaționale. Folosind metodologia elaborată, cercetarea a fost divizată în următoarele etape: identificarea problemelor, formularea ipotezelor, investigarea și analiza ipotezelor, concluzii la rezultatele obținute.

Formularea ipotezei se bazează pe scopurile și obiectivele cercetării. Una din principalele ipoteze care a influențat asupra rezultatelor finale, a fost existența extinderii pe monoidul liber a cvasimetricei ρ definite pe alfabetul A. În acest fel, unele dintre rezultatele cercetării în stadiu incipient cuprind metode eficiente de extindere a distanței pe monoizi liberi, ceea ce conduce la posibilitatea introducerii conceptului de descompunere paralelă a șirurilor de caractere.

Investigarea și analiza ipotezelor au fost efectuate în cadrul metodologiei fundamentale a cercetării științifice. La această etapă s-a stabilit că pentru orice cvasivarietate de tip non-Burnside $\mathcal{V}$ și orice cvasimetrică ρ pe o mulțime X cu punctul bază p_X pe monoidul liber $F^a(X, \mathcal{V})$ există o cvasimetrică unică stabilă $\hat{\rho}$ cu proprietățile 2.4.1 - 2.4.10, care sunt rezumate în Teorema 2.4.1 ca un rezultat principal.

Noutatea și originalitatea științifică a tezei constau în obţinerea rezultatelor noi teoretice cu aplicaţii în informatică. Rezultatele cercetărilor cuprind metode eficiente de extindere a distanțelor pe monoizi liberi, ceea ce a dus la posibilitatea introducerii conceptului de descompunere paralelă a șirurilor de caractere. Acest fapt a permis dezvoltarea conceptelor de eficiență și similaritate ale secvențelor de informație, precum și construirea mulțimilor de medii ponderate și bisectoare ale șirurilor de caractere. Gradul de noutate și originalitate este reprezentat de:

- metoda extinderii cvasimetricelor pe monoidul liber $F^a(X, \mathcal{V})$;
- studierea spațiilor digitale și Alexandroff;
- prezentarea soluțiilor problemelor lui Maltsev pentru cvasivarietățile de monoizi topologici;
- stabilirea relațiilor între distanțele Hamming, Graev și Levenshtein pe monoizi liberi;
- introducerea conceptului de eficiență a reprezentării;
- introducerea conceptului de descompuneri paralele optimale a stringurilor din monoizi liberi;
- algoritmi implementați pentru construcția mediilor ponderate și bisectoarei perechilor de șiruri de caractere;
- demonstrația ne-convexității a segmentului informațional;
- introducerea noțiunii de topologie simetrică pe linia digitală;
- demonstrația unicității topologiei Khalimsky ca topologie digitală minimală;
- elaborarea algoritmului de procesare a imaginii digitale din perspectiva topologică, aplicabilă în spațiul digital.

Problema științifică importantă rezolvată în cadrul cercetării este dezvoltarea metodelor de construire și studiere a extinderii distanțelor pe monoizi liberi, care contribuie la obținerea unor metode eficiente de reprezentare a informației, aplicabile la rezolvarea diferitor probleme de distanțe, cum ar fi alinierea secvențelor, similaritatea unei perechi de șiruri de caractere, construirea mediilor ponderate și bisectoarei unei perechi de șiruri de caractere.

Semnificația teoretică este determinată de obținerea rezultatelor noi privind stabilirea condițiilor de existență a extinderii distanțelor pe monoizi liberi ce permit construirea diverselor topologii invariante pe monoizi liberi. Metodele elaborate au permis abordarea problemelor legate de secvențele de informație dintr-un nou punct de vedere. În plus, rezultatele teoretice permit studiul liniei digitale și proprietății de minimalitate a topologiei Khalimsky.

Valoarea aplicativă a tezei constă în utilizarea rezultatelor teoretice obținute în studiul topologiilor simetrice pe linia digitală, prelucrarea imaginilor și construirea centroidului unei mulțimi de șiruri de caractere. Metodele prezentate au generat mulțimi mai mare de elemente, folosind metoda descompunerii paralele optimale.

Aprobarea rezultatelor științifice. Rezultatele științifice obținute au fost prezentate la conferințe științifice naționale și internaționale, și au fost publicate în reviste recenzate. Principalele rezultate incluse în teză au fost prezentate la următoarele conferințe științifice:

- *Scattered și Digital Topologies in Information Sciences*. Plenary talk at the Conference of the Romanian Society of Applied și Industrial Mathematics ROMAI, CAIM 2018, Chisinau, Moldova, 20-23 September 2018;
- *Scattered și 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 și 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 și Their Applications.* Conference on Applied și 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 și Their Applications*. Conference on Mathematical Foundations of Informatics, MFOI 2016, Chisinau, Moldova, 25-31 July 2016;
- *Invariant Distances on Free Semigroups și Their Applications*. The 20th Annual Conference of the Mathematical Sciences Society of Romania, 19-22 May 2016;

Publicații pe tema cercetării. Rezultatele obținute în teză sunt publicate în 20 de articole științifice (a se vedea [46]–[65]): 7 articole în reviste recenzate (a se vedea [47, 49, 56, 60, 62, 63, 64]), 13 lucrări la conferințe internaționale (a se vedea [46, 48, 50, 51, 52, 54, 53, 55, 57, 58, 59, 61, 65]); 8 publicații de un singur autor (a se vedea[46]–[53]), printre care 2 articole în reviste științifice recenzate (a se vedea [47, 49]). Volumul total al publicațiilor este de 6.4 coli de autor.

Volumul și structura tezei. Teza este scrisă în limba engleză și conține introducere, patru capitole, concluzii generale și recomandări, adnotări în limbile română, rusă și engleză, bibliografie ce cuprinde 200 de titluri. Volumul total al tezei este de 130 de pagini, dintre care 114 pagini text de bază.

3. SINTEZA CAPITOLELOR

În introducere sunt formulate actualitatea și importanța temei de cercetare. În plus, sunt definite obiectivele de cercetare, scopul, noutatea științifică și originalitatea tezei. Problema științifică studiată este prezentată cu accent pe importanța valorii teoretice și aplicative a lucrării. Se oferă o scurtă analiză a problemelor și a publicațiilor pe tema tezei. Această secțiune se încheie cu un rezumat al conținutului lucrării.

Primul capitol, Situația actuală în domeniul teoriei spațiului cvasimetric și aplicațiile lor în algebra și teoria informației , are un caracter introductiv și conține un rezumat al celor mai importante rezultate conform scopului si obiectivelor tezei, direcțiilor de cercetare. Se definesc și se clasifică distanțe, spații de distanță, sisteme informatice de tip Scott-Ershov, algebre topologice universale, spatii de șiruri de caractere.

Pentru rezolvarea problemei de cercetare definite este importantă studierea următoarelor probleme particulare:

1. Determinarea condițiilor de extindere a cvasimetricei ρ pe A și orice cvasivarietăți $\mathcal{V}$ la o cvasimetrică invariantă ρ^* pe monoidul liber $F^a(A, \mathcal{V})$.
2. Propunerea algoritmilor de calcul al distanței $\rho^*(a, b)$ între două secvențe de informație $a, b \in L(A)$.
3. Determinarea relațiilor între proprietățile topologo-geometrice a spațiilor (A, ρ) și $(L(A), \rho^*)$.
4. Propunerea metodelor de construcție a mulțimilor de medii ponderate și bisectoarei ale unei perechi de șiruri date.
5. Determinarea proprietăților topologo-geometrice care sunt importante în analiza informației și a procesării imaginilor.

Problemele lui Maltsev sunt formulate pentru monoizi liberi, care au fost puse în 1957 [33]:

Problema Maltsev I: În ce condiții aplicația v_X este o scufundare?

Problema Maltsev II: În ce condiții homomorfismul w_X este un izomorfism continuu?

Pentru spațiile complet regulate X, problemele lui Maltsev au fost rezolvate afirmativ de S. Swierczkowski [42] în cazul semnăturii discrete E, și de către M. M. Choban și S. S. Dumitrashcu pentru orice semnătura [19, 14].

La sfârșitul acestui capitol, se formulează problema cercetării, se identifică metodele de rezolvare a acesteia, se stabilesc obiectivele și scopurile cercetării.

În capitolul doi, Extinderea cvasimetricelor pe monoizi liberi topologici, sunt studiate cvasivarietățile de monoizi topologici și sunt elaborate noi metode de extindere a cvasimetricelor. Capitolul începe cu introducerea monoizilor topologici liberi și construirea monoidului liber abstract.

Orice spațiu topologic X este considerat a fi un spațiu Kolmogorov, adică un T_0-spațiu: pentru oricare două puncte distincte $x, y \in X$ există o submulțime deschisă U din X astfel încât $U \cap \{x, y\}$

este o mulțime singleton. Clasa $\mathcal{V}$ de monoizi topologici se numește cvasivarietate de monoizi dacă: (1) clasa $\mathcal{V}$ este multiplicativă; (2) dacă $G \in \mathcal{V}$ și A este un submonoid din G, atunci $A \in \mathcal{V}$; (3) fiecare spațiu $G \in \mathcal{V}$ este un T_0-spațiu. Un monoid liber al unui spațiu X într-o clasă $\mathcal{V}$ este un monoid topologic $F(X, \mathcal{V})$ cu proprietățile: $X \subseteq F(X, \mathcal{V}) \in \mathcal{V}$ și p_X este unitatea lui $F(X, \mathcal{V})$, mulțimea X generează monoidul $F(X, \mathcal{V})$ și pentru orice aplicație continuă $f : X \longrightarrow G \in \mathcal{V}$, unde $f(p_X) = e$, există un homomorfism unic continuu $\bar{f} : F(X, \mathcal{V}) \longrightarrow G$ astfel încât $f = \bar{f}|X$. Un monoid abstract liber al unui spațiu X într-o clasă $\mathcal{V}$ este un monoid topologic $F^a(X, \mathcal{V})$ cu proprietățile: X este o submulțime din $F^a(X, \mathcal{V})$, $F^a(X, \mathcal{V}) \in \mathcal{V}$ și p_X este unitatea lui $F^a(X, \mathcal{V})$, mulțimea X generează monoidul $F^a(X, \mathcal{V})$, iar pentru orice aplicație $f : X \longrightarrow G \in \mathcal{V}$, unde $f(p_X)$ $= e$, există un homomorfism continuu unic $\hat{f} : F^a(X, \mathcal{V}) \longrightarrow G$ astfel încât $f = \hat{f}|X$.

Prin metoda Kakutani se demonstrează că:

Teorema 2.1.1. *Fie $\mathcal{V}$ o cvasivarietate netrivială de monoizi topologici. Atunci pentru fiecare spațiu X următoarele afirmații sunt echivalente:*

1. *Există $G \in \mathcal{V}$ astfel încât X este un subspațiu din G și p_X este elementul neutru în G.*
2. *Pentru spațiul X există un unic monoid liber topologic $F(X, \mathcal{V})$.*

Continuăm cu evidențierea problemelor abordate în această secțiune.

Problema 2.1.1. *Fie $\mathcal{V}$ o cvasivarietate netrivială de monoizi topologici. În ce condiții pentru un spațiu X există monoidul liber topologic $F(X, \mathcal{V})$?*

Fixăm un spațiu X pentru care există monoidul liber topologic $F(X, \mathcal{V})$. Atunci există un homomorfism unic continuu $\pi_X : F^a(X, \mathcal{V}) \longrightarrow F(X, \mathcal{V})$ astfel încât $\pi_X(x) = x$ pentru orice $x \in X$. Monoidul $F(X, \mathcal{V})$ se numește abstract liber dacă π_X este un izomorfism continuu.

Problema 2.1.2. *Fie $\mathcal{V}$ o cvasivarietate netrivială de monoizi topologici. În ce condiții pentru un spațiu X există monoidul liber topologic $F(X, \mathcal{V})$, care este abstract liber?*

Problemele 2.1.1 și 2.1.2 sunt importante în teoria algebrelor universale cu topologii (a se vedea [33, 14, 15, 12, 13, 16]). Aceste probleme pentru varietăți de algebre topologice au fost formulate de A. I. Maltsev ([33], vezi problemele lui Maltsev în secțiunea 1.6 a tezei).

Următoarea teoremă se bazează pe rezultatul menționat în teorema anterioară 2.1.1.

Teorema 2.1.2. *Fie $\mathcal{V}$ o cvasivarietate netrivială de monoizi topologici, și fie că există $H \in \mathcal{V}$ și punctul $b \in H$, astfel încât $e \neq b$ și $E = \{e, b\}$ este un subspațiu discret din H. Atunci pentru fiecare spațiu zero-dimensional X există monoidul liber topologic unic $F(X, \mathcal{V})$.*

Exemplul 2.1.1 ilustrează faptul că nu pentru orice cvasivarietate netrivială $\mathcal{V}$ și orice T_0-spațiu X există $F(X, \mathcal{V})$.

Fie $\omega = \{0, 1, 2, \ldots\}$. O cvasivarietate $\mathcal{V}$ de monoizi topologici se numește cvasivarietate Burnside dacă există două numere minimale $p = p(\mathcal{V})$, $q = q(\mathcal{V}) \in \omega$, astfel încât $0 \leq q < p$ și $x^p = x^q$ pentru toate $x, y \in G \in \mathcal{V}$. În acest caz, orice $G \in \mathcal{V}$ este un monoid (p, q)-periodic cu exponentul (p, q). Dacă $q = 0$, atunci orice monoid $G \in \mathcal{V}$ este un monoid periodic cu exponentul p

și $x^p = e$ pentru fiecare $x \in G \in \mathcal{V}$. Cvasivarietatea trivială este considerată Burnside cu exponentul $(0, 1)$.

Următoarea teoremă rezolvă problema 2.1.1 pentru cvasivarietățile non-Burnside complete de monoizi topologici .

Teorema 2.3.3. *Fie $\mathcal{V}$ o cvasivarietate completă non-Burnside de monoizi topologici. Atunci pentru fiecare T_0-spațiu X există monoidul liber topologic $F(X, \mathcal{V})$.*

Următoarea teoremă rezolvă problema 2.1.1 pentru cvasivarietății complete netriviale de monoizi topologici.

Teorema 2.3.4. *Fie $\mathcal{V}$ o cvasivarietate completă netrivială de monoizi topologici. Atunci pentru fiecare spațiu complet regulat X există monoidul liber topologic $F(X, \mathcal{V})$.*

Fixăm o cvasivarietate completă netrivială $\mathcal{V}$ de monoizi topologici. Considerăm o mulțime nevidă X cu un punct fix $e \in X$. Presupunem că $e \in X \subseteq F^a(X, \mathcal{V})$ și e este unitatea monoidului $F^a(X, \mathcal{V})$. Fie ρ o pseudo-cvasimetrică pe mulțimea X. Notăm cu $Q(\rho)$ mulțimea tuturor pseudo-cvasimetricilor stabile d pe $F^a(X, \mathcal{V})$ pentru care $d(x, y) \leq \rho(x, y)$, pentru toate $x, y \in X$. Mulțimea $Q(\rho)$ este nevidă pentru că ea conține pseudo-cvasimetrica trivială $d(x, y) = 0$, pentru toate $x, y \in F^a(X, \mathcal{V})$. Pentru orice $a, b \in F^a(X, \mathcal{V})$ punem $\hat{\rho}(a, b) = sup\{d(a, b) : d \in Q(\rho)\}$. Spunem că $\hat{\rho}$ este extinderea maximală stabilă a ρ pe $F^a(X, \mathcal{V})$. Pentru orice $a, b \in F^a(X, \mathcal{V})$ punem $\bar{\rho} = inf\{\Sigma\{\rho(x_i, y_i) : i \leq n\} : n \in \mathbb{N}, x_1, y_1, x_2, y_2, ..., x_n, y_n \in X, a = x_1x_2...x_n, b = y_1y_2...y_n\}$ și $\rho^*(a, b) = inf\{\bar{\rho}(a, z_1) + ... + \bar{\rho}(z_i, z_{i+1}) + ... + \bar{\rho}(z_n, b) : n \in \mathbb{N}, z_1, z_2, ..., z_n \in F^a(X, \mathcal{V})\}$.

Unul dintre rezultatele principale ale tezei este următoarea teoremă.

Teorema 2.4.1. *Fie ρ o pseudo-cvasimetrică pe X, Y un subspațiu al lui X și $e \in Y$. Notăm cu $M(Y) = F^a(Y, \mathcal{V})$ submonoidul monoidului $F^a(X, \mathcal{V})$ generat de mulțimea Y și cu d_Y extinderea $\hat{\rho}|Y$ pe $M(Y)$ a pseudo-cvasimetricei ρ_Y pe Y, unde $\rho_Y(y, z) = \rho(y, z)$ pentru toate $y, z \in Y$. Atunci:*

1. *$d_Y(a, b) = \hat{\rho}(a, b)$ pentru toate $a, b \in M(Y)$.*
2. *Dacă $\mathcal{V}$ este o cvasivarietate non-Burnside, atunci $\bar{\rho}(x, y) = \rho(x, y)$ pentru toate $x, y \in X$.*
3. *Dacă ρ este o cvasimetrică (strictă) pe Y, atunci $\hat{\rho}$ este o cvasimetrică (strictă) pe $M(Y)$.*
4. *Dacă ρ este o a metrică pe Y, atunci $\hat{\rho}$ este o metrică pe $M(Y)$.*
5. *Dacă $a, b \in F^a(Y, \mathcal{V})$ sunt puncte distincte și ρ este o cvasimetrică pe $Sup(a, b)$, atunci $\hat{\rho}(a, b) + \hat{\rho}(b, a) > 0$.*
6. *Dacă $a, b \in F^a(Y, \mathcal{V})$ sunt puncte distincte și ρ este o cvasimetrică strictă pe $Sup(a, b)$, atunci $\hat{\rho}(a, b) > 0$ și $\hat{\rho}(b, a) > 0$.*
7. *Pentru orice $a, b \in F^a(Y, \mathcal{V})$ există $n \in \mathbb{N}$, $x_1, x_2, ..., x_n \in Sup(a, a)$ și $y_1, y_2, ..., y_n \in Sup(b, b)$ astfel încât $a = x_1x_2...x_n$, $b = y_1y_2...y_n$, $n \leq l(a) + l(b)$ și $\bar{\rho}(a, b) = \Sigma\{\rho(x_i, y_i) : i \leq n\}$.*

8. $\hat{\rho} = \bar{\rho} = \rho^*$.

În clasa monoizilor liberi, Teorema 2.4.1 permite rezolvarea celor doua probleme ale lui Maltsev, formulate pentru algebre universale în 1957 [33].

Cvasivarietatea de monoizi topologici $\mathcal{V}$ este rigidă dacă pentru orice spațiu X, orice cuvânt $a \in F(X, \mathcal{V})$, orice punct $c \in X \setminus \{p_x\}$ și orice reprezentare $ac = x_1x_2...x_n$, unde $x_1, x_2, ..., x_n \in X$, existsă $m \leq n$, astfel încât $x_m = c$ și $a = x_1x_2...x_{m-1}$. În acest caz $x_i = p_X = e$, pentru orice $i > m$.

Varietatea tuturor monoizilor topologici este rigidă.

Teorema 2.5.1. *Fie $\mathcal{V}$ o cvasivarietate rigidă non-Burnside de monoizi topologici, ρ o cvasimetrică pe un spațiu X cu punctul bază p_X și $\rho(x, p_X) = \rho(y, p_X)$, pentru orice $x, y \in X \setminus \{p_X\}$, sau $\rho(p_X, x)$ $= \rho(p_X, y)$, pentru orice $x, y \in X \setminus \{p_X\}$. Atunci $\rho^*(ac, bc) = \rho^*(ca, cb) = \rho^*(a, b)$, pentru toate $a, b, c \in F(X, \mathcal{V})$.*

Următoarea teoremă extinde Teorema 2.3.3 și rezolvă problema 2.1.2 pentru cvasivarietăți complete non-Burnside de monoizi topologici.

Teorema 2.6.1. *Fie $\mathcal{V}$ o cvasivarietate netrivială completă non-Burnside de monoizi topologici. Atunci:*

1. Pentru orice T_0-spațiu X pe monoidul liber $F^a(X, \mathcal{V})$ există o T_0-topologie $\mathcal{T}(qm)$, astfel încât:

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

– X este un subspațiu al spațiului $(F^a(X, \mathcal{V}), \mathcal{T}(qm))$;

– topologia $\mathcal{T}(qm)$ este generată de familia tuturor pseudo-cvasimetricilor invariante pe $F^a(X, \mathcal{V})$, care sunt continue pe X.

2. Pentru orice T_0-spațiu X monoidul liber topologic $F(X, \mathcal{V})$ există și este abstract liber.

3. Un spațiu X este un T_1-spațiu dacă si numai dacă spațiile $F(X, \mathcal{V})$ și $(F^a(X, \mathcal{V}), \mathcal{T}(qm))$ sunt T_1-spații.

4. Un spațiu X este funcțional Hausdorff dacă si numai dacă spațiile $F(X, \mathcal{V})$ și $(F^a(X, \mathcal{V}), \mathcal{T}(qm))$ sunt funcțional Hausdorff.

Rezultatele analogice cu Teoremele 2.1.1 și 2.6.1 sunt obținute pentru monoidul semi-topologic $F(X, \mathcal{W})$ în Teoremele 2.7.1 și 2.7.2.

Al doilea capitol se încheie cu discuția despre spațiile digitale topologice, rezultatele rezumate în Corolarul 2.8.2, care rezultă din Corolarul 2.8.1 și Propozițiile 2.8.2 și 2.8.4.

Corolarul 2.8.1. *Fie $\mathcal{V}$ o cvasivarietate netrivială completă non-Burnside de monoizi topologici. Atunci pentru fiecare spațiu X următoarele afirmații sunt echivalente:*

1. $F(X, \mathcal{V})$ este un spațiu Alexandroff.

2. Pe spațiul $F(X, \mathcal{V})$ există o cvasimetrică cu valori naturale.

3. X este un spațiu Alexandroff.

Corolarul 2.8.2. *Fie $\mathcal{V}$ o cvasivarietate netrivială completă non-Burnside de monoizi topologici. Atunci pentru fiecare spațiu X următoarele afirmații sunt echivalente:*

1. $F(X, \mathcal{V})$ este spațiu topologic digital.

2. X este spațiu topologic digital.

Rezultatele prezentate în acest capitol completează cu succes lucrările matematicienilor în domeniul extinderii distanțelor pe structurile algebrice abstracte. Cercetarea autorului în acest capitol este publicată în articolele [54, 57, 62] și servește drept bază pentru cercetarea prezentată în următoarele capitole.

În capitolul 3, Măsura similarității pe monoizi de stringuri, se demonstrează că există distanțe invariante pe monoidul $L(A)$ al tuturor stringurilor asemanatoare cu distanțele Levenshtein și Hamming. Se introduce o definiție distinctă a distanței pe $L(A)$, pe baza metodei Markov-Graev, propusă pentru grupuri libere. În rezultat, se arată că pentru orice cvasimetrică d pe alfabetul A în reuniune cu șirul vid există o extendere maximală invariantă d^* pe monoidul liber $L(A)$. Această nouă abordare permite introducerea descompunerilor paralele și semiparalele a două stringuri, care sunt ulterior utilizate în subsecțiunea 3.3 pentru a determina factorii de eficiență și de penalizare în două reprezentări de stringuri.

Pentru orice stringuri $a, b \in L(A)$ găsim descompuneri de forma $a = v_1u_1v_2u_2 \ldots v_ku_kv_{k+1}$ și $b = w_1u_1w_2u_2 \ldots w_ku_kw_{k+1}$, care pot fi reprezentate ca $a = a_1a_2 \ldots a_n$, $b = b_1b_2 \ldots b_n$ cu următoarele proprietăți:

- unele a_i și b_j pot fi stringuri vide, adică $a_i = \varepsilon$, $b_j = \varepsilon$;

- dacă $a_i = \varepsilon$, atunci $b_i \neq \varepsilon$, și dacă $b_j = \varepsilon$, atunci $a_j \neq \varepsilon$;

- dacă $u_1 = \varepsilon$, atunci $a = v_1$ și $b = w_1$;

- dacă $u_1 \neq \varepsilon$, atunci există șirul $1 \leq i_1 \leq j_1 < i_2 \leq j_2 < \ldots < i_k \leq j_k \leq n$, astfel încât: $u_1 = a_{i_1} \ldots a_{j_1} = b_{i_1} \ldots b_{j_1}, u_2 = a_{i_2} \ldots a_{j_2} = b_{i_2} \ldots b_{j_2}, u_k = a_{i_k} \ldots a_{j_k} = b_{i_k} \ldots b_{j_k}$;

- dacă $v_1 = w_1 = \varepsilon$, atunci $i_1 = 1$;

- dacă $v_{k+1} = w_{k+1} = \varepsilon$, atunci $j_k = n$;

- dacă $k \geq 2$, atunci pentru orice $i \in \{2, \ldots, k\}$ avem $v_i \neq \varepsilon$ sau $w_i \neq \varepsilon$.

În acest caz avem $l(u_1) + l(u_2) + \ldots + l(u_k) = |\{i : a_i = b_i\}|$.

Descompunerile care au forma de mai sus se numesc *descompuneri paralele* a stringurilor a și b [55, 56, 57]. Pentru orice descompuneri paralele $a = v_1u_1 \ldots v_ku_kv_{k+1}$ și $b = w_1u_1 \ldots w_ku_kw_{k+1}$ numărul

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

se numește eficiența descompunerilor paralele date. Numărul $E(a, b)$ este egal cu valoarea minimă a eficienței tuturor descompunerilor paralele ale stringurilor a, b și se numește *eficiența comună a stringurilor a, b*. E evident că $E(a, b)$ este bine determinată și $E(a, b) = d_G(a, b)$. Spunem că descompunerile paralele $a = v_1u_1v_2u_2 \ldots v_ku_kv_{k+1}$ și $b = w_1u_1w_2u_2 \ldots w_ku_kw_{k+1}$ sunt optimale dacă are loc următoarea egalitate:

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

Acest tip de descompuneri sunt asociate cu problema asemănării aproximative a șirurilor [35]. Dacă descompunerile $a = v_1u_1 \ldots v_ku_kv_{k+1}$ și $b = w_1u_1 \ldots w_ku_kw_{k+1}$ sunt optimale și $k \geq 2$, atunci putem considera că $u_i \neq \varepsilon$ pentru orice $i \leq k$.

Orice descompuneri paralele $a = a_1a_2 \ldots a_n = v_1u_1v_2u_2 \ldots v_ku_kv_{k+1}$ și $b = b_1b_2 \ldots b_n = w_1u_1w_2u_2 \ldots w_ku_kw_{k+1}$ generează subșiruri comune $u_1u_2 \ldots u_k$. Numărul

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

se numește *măsura similarității* a descompunerilor [7, 37]. Există descompuneri paralele $a = v_1u_1v_2u_2\ ldots v_ku_kv_{k+1}$ și $b = w_1u_1w_2u_2\ ldots w_ku_kw_{k+1}$ pentru care măsura de similaritate este maximă. Valoarea maximă a măsurii de similaritate a tuturor descompunerilor se notează cu $m^*(a, b)$. Valoarea maximă a măsurii de similaritate a tuturor descompunerilor optimale se notează cu $m^\omega(a, b)$. Observăm că $m^\omega(a, b) \leq m^*(a, b)$. Pentru orice două descompuneri paralele $a = a_1a_2 \ldots a_n$ și $b = b_1b_2 \ldots b_n$ prin analogie cu [56], definim *factorul de penalitate*

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

și *măsurile de similaritate proprie*

$$\begin{aligned}
M_r(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) &= m(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) - p_r(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n), \\
M_l(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) &= m(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) - p_l(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n), \\
M(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) &= m(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) - p(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n)
\end{aligned}$$

.

Numărul $d_H(a_1a_2 \ldots a_n, b_1b_2 \ldots b_n) = |\{i \leq n : a_i \neq b_i\}|$ este distanța Hamming dintre discompuneri și este un alt tip de penalitate. Avem:

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

Afirmațiile din următoarea teoremă stabilesc principalele rezultate.

Teorema 3.3.1. *Fie a și b două stringuri nevide, $a = a_1a_2 \ldots a_n$ și $b = b_1b_2 \ldots b_n$ descompuneri optimale inițiale, iar $a = a'_1a'_2 \ldots a'_q$ și $b = b'_1b'_2 \ldots b'_q$ a doua pereche de descompuneri arbitrare. Notăm:*

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

$M = m - p$, $M' = m' - p'$, $M_l = m - p_l$, $M'_l = m' - p'_l$, $M_r = m - p_r$, $M'_r = m' - p'_r$.
Următoarele afirmații sunt adevărate:

1. $p' - p = 2(m' - m) + 2(r' - r)$.
2. *Dacă a doua pereche de descompuneri nu este optimală, atunci* $M_l > M'_l$ *și* $M_r > M'_r$.
3. *Dacă a doua pereche de descompuneri este optimală, atunci* $M_l = M'_l$ *și* $M_r = M'_r$, *și măsurile* M_l *și* M_r *sunt constante pe mulțimea descompunerilor paralele optimale.*
4. *Dacă* $m' \geq m$ *și a doua pereche de descompuneri nu este optimală, atunci* $p' > p$, $p_{l'} > p_l$, $p'_r > p_r$ *și* $M > M'$.
5. *Dacă* $m' = m$ *și a doua pereche de descompuneri este optimală, atunci* $p' = p$, $p_{l'} = p_l$, $p'_r = p_r$ *și* $M' = M$.
6. *Dacă* $m' \leq m$ *și a doua pereche de descompuneri nu este optimală, atunci* $m' - r' < m - r$.

Din afirmațiile teoremei 3.3.1 rezultă că pe clasa tuturor descompunerilor optimale a două stringuri date:

- măsura maximală de similaritate proprie se realizează pe descompunerea paralelă optimală cu penalitate minimală (măsura minimă de similaritate);
- măsura minimă de similaritate proprie se realizează pe descompunerea paralelă optimală cu penalitatea maximă (măsura maximă de similaritate).

Pentru oricare două șiruri nevide există descompuneri paralele cu măsura maximă de similaritate și descompuneri optimale pe care măsura de similaritate este minimă.

Prezentăm mai jos un exemplu care ilustrează relațiile privind similaritate proprie și penalitățile impuse de afirmația 4 a teoremei 3.3.1.

Exemplul 3.3.1. *Fie* $a = AAAACCC$, $b = CCCBBBB$ *două stringuri cu a și b fiind descompunerile lor optimale triviale, și* $a' = AAAACCC\varepsilon\varepsilon\varepsilon\varepsilon$, $b' = \varepsilon\varepsilon\varepsilon\varepsilon CCCBBBB$ *ca descompunerile lor neoptimale. Atunci* $m' = 3$, $r' = 8$, $p' = 8$,$m = 0$, $r = 7$, *și* $p = 0$. *În acest exemplu avem* $-5 = m' - r' > m - r = -7$ *și* $-5 = m' - p' = M' < M = m - p = 0$.

Următorul exemplu arată că există unele descompuneri exotice neoptimale paralele $a = a'_1a'_2\cdots a'_q$ și $b = b'_1b'_2\cdots b'_q$, astfel încât pentru descompunerile optimale $a = a_1a_2\cdots a_n$ și $b = b_1b_2\cdots b_n$ avem $m' < m$, $p' < p$, și $M' > M$.

Exemplul 3.3.2. *Fie $ABCDEF$ și $CDEFED$ descompunerile triviale neoptimale a stringurilor a, b și $ABCDEF\varepsilon\varepsilon$, $\varepsilon\varepsilon CDEFED$ descompunerile lor optimale. Atunci $m' = 1$, $r' = 5$, $p' = 0$, $m = 4$, $r = 4$, și $p = 4$. Obţinem că $m' - p' = M' > M = m - p$, și $m' - r' < m - r$.*

Exemplele de mai sus arată că Teorema 3.3.1 nu poate fi îmbunătățită în cazul $m' < m$.

Descompuneri cu penalitatea minimă și similaritatea proprie maximală sunt de interes semnificativ. Mai mult, dacă rezolvăm problema de editare și corectare a textelor, descompunerile optimale sunt mai favorabile. Prin urmare, descompunerile optimale sunt cele mai bune descompuneri paralele și putem rezolva probleme de aliniere a stringurilor numai pe clasa de descompuneri optimale.

În virtutea Teoremei 3.3.1, diferite aplicații ale distanțelor pe monoizi de șiruri pot fi folosite în rezolvarea problemelor din diferite domenii științifice. De exemplu, studiul măsurii de similaritate proprie este abordat dintr-o perspectivă nouă.

Remarca 3.5.1. *Algoritmii noștri sunt eficienți pentru orice cvasimetrică pe $\bar{A}$. Unii autori consideră posibilitatea de a defini metrica generalizată Levenshtein cu valori distincte $\rho(a, b)$ și $\rho(b, a)$. Este necesar ca $\rho(a, b)$ să fie o cvasimetrică. În caz contrar, putem obține confuzii după cum se va vedea din exemplul următor.*

Exemplul 3.5.4. *Fie $A = \{a, b\}$, $\bar{A} = \{\varepsilon, a, b\}$. Următorul tabel definește distanța ρ pe $\bar{A}$:*

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

În acest exemplu avem $0 = \rho(a, b) + \rho(b, \varepsilon) < \rho(a, \varepsilon) = 1$ și:

1. *pentru $u = aba$, $v = ba$ obţinem $\bar{\rho}(u, v) = \bar{\rho}(v, u) = 0$,*
2. *pentru $u = a$, $v = b$ obţinem $\bar{\rho}(u, v) = \bar{\rho}(v, u) = 0$, iar $\rho(v, u) = 1$.*

Exemplul 3.5.5. *Să examinăm exemplul din [37] în contextul rezultatelor obținute. Avem stringurile $a = AJCJNRCKCRBP$ și $b = ABCNJROCLCRPM$ pentru care sunt opt perechi de descompuneri paralele optimale. Prezentăm două din ele, cea mai scurtă și cea mai lungă:*

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

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

Pentru prima pereche avem $\rho^* = 7$, $m = 6$, $p = 1$, și $M = 5$. Pentru a doua pereche avem $\rho^* = 7$, $m = 8$, $p = 5$, și $M = 3$. Algoritmii noștri ne permit să calculăm toate descompunerile optimale cu măsuri de similaritate distincte. Autorii din [37] preferă cea de-a doua pereche de descompuneri pentru că are o măsură maximală de similaritate. Considerăm mai preferabilă prima pereche, care are similaritatea proprie maximală.

De fapt, pentru orice două șiruri nevide există descompuneri paralele cu o măsură maximă de similaritate și descompuneri optimale pe care măsură de similaritate este minimă. Exemplul următor arată că există unele descompuneri exotice paralele neoptimale $a = a'_1 a'_2 \cdots a'_q$ și $b = b'_1 b'_2 \cdots b'_q$, astfel încât pentru descompunerile optimale $a = a_1 a_2 \cdots a_n$ și $b = b_1 b_2 \cdots b_n$ avem $m' < m$, $p' < p$ și $M' > M$.

Exemplul 3.5.6. *Fie $a = ABCDEF$ și $b = CDEFED$ descompunerile triviale neoptimale ale stringurilor a, b, cu $a = ABCDEF\varepsilon\varepsilon$ și $b = \varepsilon\varepsilon CDEFED$ descompunerile lor optimale. Atunci $m' = 1$, $r' = 5$, $p' = p'_l = p'_r = 0$ și $m = 4$, $r = 4$, $p = 4$, $p_l = p_r = 2$. În acest exemplu avem $M'_l = M'_r = M' = m' - p' = 1 - 0 = 1 > 0 = 4 - 4 = m - p = M$, $m' - r' = -4 < 0 = m - r$, $M_l = 4 - 2 = 2 > 1 = M'_l$, $M_r = 4 - 2 = 2 > 1 = M'_r$.*

Exemplul 3.5.7. *Fie $a = AAAACCC$ și $b = CCCBBBB$ descompunerile triviale optimale ale stringurilor a, b, cu $a = AAAACCC\varepsilon\varepsilon\varepsilon\varepsilon$ și $b = \varepsilon\varepsilon\varepsilon\varepsilon CCCBBBB$ descompunerile lor neoptimale. Atunci $m' = 3$, $r' = 8$, $p' = 8$ și $m = 0$, $r = 7$, $p = p_l = p_r = 0$. În acest exemplu avem $-5 = m' - r' > m - r = -7$ și $-5 = m' - p' < m - p = 0$.*

Exemplele de mai sus arată că Teorema 3.3.1 nu poate fi îmbunătățită în cazul $m' < m$. Proprietățile examinate și rezultatele obținute în acest capitol sunt publicate în articolele [46, 47, 51, 55, 56, 57, 62, 63] și servesc drept bază pentru capitolul următor. Rezultatele menționate pot fi aplicate în diferite probleme legate de similaritatea între secvențe de caractere.

Capitolul 4, Aspectele geometrice si topologice ale analizei informației, este capitolul final al tezei și se concentrează pe partea aplicativă a rezultatelor teoretice obținute în capitolul precedent. Mai precis, în acest capitol este rezolvată problema construirii mulțimelor de medii ponderate și bisectoarei perechelor de șiruri de caractere. Acest rezultat se bazează puternic pe proprietățile descompunerilor paralele optimale. Se demonstrează că orice element al mulțimii de medii ponderate este generat de unele descompuneri paralele optimale. Această proprietate puternică, precum și reciproca ei, sunt rezumate în următoarele două teoreme.

Teorema 4.1.1. *Orice descompunere paralelă fixată d - optimală a unei perechi date de stringuri $a, b \in L(A)$ generează mediile ponderate, simultan cu reprezentările lor echivalente, care formează descompuneri paralele d - optimale cu reprezentările fixate ale stringurilor date.*

Corolarul 4.1.1. *Orice medie ponderată a unei perechi fixate de stringuri este generată de careva din descompunerile sale paralele optimale.*

Ca un caz special, atunci când d_G (distanța Graev) este o metrică discretă și $d_G(a, b)$ este un număr par, media ponderată egal îndepărtat de stringurile a și b este mediana segmentului determinat de a și b. Algoritmul pentru construirea medianelor unui șir de perechi este prezentat în Algoritmul 5.

Algorithm 5: Mediana DescParalOptim a x și y:
Date $x, y \in L(\bar{A})$ construiește $m \in L(\bar{A})$, cu $d^*(x, m) = d^*(m, y)$.

Data: $x = x_1 x_2 \dots x_n$, $y = y_1 y_2 \dots y_m$.
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**
    $m := m_1 m_2 \dots m_{|I|}$, 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;

Capitolul continuă cu studiul problemei convexității mulțimii de medii ponderate. Teoremele care urmează, abordează aceste întrebări utilizînd distanțele Hamming și Graev.

Întrebarea 1. *Este oare mulțimea $M_{d_H}(a, b)$ d_H-convexă în $(L^*(A), d_H)$ pentru orice $a, b \in L^*(A)$?*

Întrebarea 2. *Este oare mulțimea $M_{d_G}(a, b)$ d_G-convexă în $(L^*(A), d_G)$ pentru orice $a, b \in L^*(A)$?*

Teorema 4.2.1. *Mulțimea $M_{d_H}(a, b)$ este d_H-convexă în $(L^*(A), d_H)$ pentru orice $a, b \in L^*(A)$.*

Teorema 4.2.2. *Există un alfabet finit A și două stringuri $a, b \in L(A)$ pentru care mulțimea $M_{d_G}(a, b)$ nu este d_G-convexă.*

Următorul exemplu prezintă stringurile $a, b \in L(A)$ pentru care $M_{d_G}(a, b)$ nu este d_G-convexă.

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

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

Considerăm stringurile amb, bma, $a'ma'$, $b'mb'$ și cmc. Obținem:

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

Deci $a'ma', b'mb'$ sunt din mijlocul segmentului $M_{d_G}(amb, bma)$, stringul cmc este din mijlocul segmentului $M_{d_G}(a'ma', b'mb')$, iar $cmc \notin M_{d_G}(amb, bma)$.

Capitolul continuă cu teorema cu privire la metodele de construcție a elementelor din mulțimea bisectoarei a două stringuri, adică șiruri care sunt egal depărtate de la a și b.

Teorema 4.3.1. *Fie $a = a_1a_2...a_n$ și $b = b_1b_2...b_n$ două stringuri din $L^*(A)$. Există metode de construcție a elementelor $c = c_1c_2...c_n \in \bar{B}_{d_H}(a, b)$.*

Capitolul se încheie cu studiul metodelor de prelucrare a imaginilor folosind noțiunile de spații digitale dispersate. Unul dintre rezultatele acestui studiu stabilește că topologia Khalimsky este topologia digitală minimă din clasa tuturor topologiilor simetrice de pe linia discretă $\mathbb{Z}$. Spunem că topologia $\mathcal{T}$ pe $\mathbb{Z}$ este simetrică dacă $(\mathbb{Z}, \mathcal{T})$ este un spatiu Alexandroff dispersat, mulțimea $\{0\}$ nu este deschisă în $(\mathbb{Z}, \mathcal{T})$ și pentru orice $n \in \mathbb{Z}$ aplicația $S_n : \mathbb{Z} \to \mathbb{Z}$, unde $S_n(x) = 2n - x$ pentru toate $x \in \mathbb{Z}$, este un homeomorfism.

Este bine cunoscut faptul că diferite structuri topologice algebrice au fost introduse pentru a răspunde cerințelor teoriei informaționale. În procesul de studiere a obiectelor continue prin metode computerizate, ele sunt aproximate prin obiecte finite sau prin imagini digitale [1, 9, 11, 56, 30, 38, 39, 41].

Procesarea digitală a imaginilor este un proces care, din punct de vedere topologic, poate fi descris în felul următor:

1. Fixăm un spațiu infinit X (o imagine continuă a originalului) și o proprietate $\mathcal{P}$ de subspații ale spațiului X.

2. Prin careva procedură vom construi un număr $n \in \omega$, o submulțime finită $H = \{h_i : i \in \omega(n)\} \subset \mathbb{Z}$ de nivele și o familie finită $\{G_i : i \in \omega(n)\}$ de submulțimi nevide deschise a spațiului X cu proprietățile:

- $G_i \cap G_k = \emptyset$ pentru orice $0 \leq i < k \leq n$;

- pentru orice $i \in \omega(n)$ și orice $x \in G_i$ există o submulțime deschisă $G(x)$ astfel încât $x \in G(x) \subset G_i$ și $G(x)$ este o submulțime cu proprietatea $\mathcal{P}$ în X;

- mulțimea $G = \{G_i : i \in \omega(n)\}$ este densă în X.

Mulțimea G este $\mathcal{P}$-nucleu și $X \setminus G$ este $\mathcal{P}$-reziduu al spațiului X.

3. Aplicația intensității $I_{\mathcal{P}} : X \to \omega \subseteq H$ este determinată cu properietatea: $I_{\mathcal{P}}(x) = maximal\{h_i : x \in cl_X G_i\}$ pentru orice $x \in X$. Avem $G_i \subset I_{\mathcal{P}}^{-1}(h_i)$ pentru orice $i \in \omega(n)$.

4. Pe H se determină topologia digitală pentru care aplicația $I_{\mathcal{P}}$ este continuă.

5. Printr-o procedură se construiește un T_0-spațiu finit K și pentru orice $x \in X$ se determină o submulțime nevidă $D_{\mathcal{P}}(x)$ din K, astfel încât:

- pentru orice $c \in K$ mulțimea $X(c) = \{x \in X : c \in D_{\mathcal{P}}(x)\}$ este închisă și se numește o $\mathcal{P}$-celulă a X;

- pentru orice $c \in K$ există $i \in \omega(n)$ și o submulțime deschisă nevidă $X'(c) \subset G_i$, astfel încât $X(c) = cl_X X'(c)$.

Rezultatele cercetărilor prezentate în acest capitol, împreună cu rezultatele discutate în capitolele anterioare, acoperă pe deplin obiectivele de cercetare menționate în primul capitol. Rezultatele din acest capitol sunt publicate în articolele [49, 50, 60, 61] și pot fi aplicate în studiul unui cerc larg de probleme teoretice și practice.

4. CONCLUZII GENERALE

Cercetările efectuate în cadrul tezei "Distanțe pe Monoizi Liberi și Aplicațiile lor în Teoria Informației" corespund în întregime scopului și obiectivelor expuse în întroducerea lucrării.

Studiul rezultatelor obținute permite evidențierea următoarelor rezultate generale:

1. S-a stabilit că pentru orice cvasivarietate non-Burnside $\mathcal{V}$ și orice cvasimetrică ρ pe o mulțime X cu punctul bază p_X pe monoid liber $F^a(X, \mathcal{V})$ există o cvasimetrică unică stabilă $\hat{\rho}$ cu proprietățile:
   (a) $\rho(x, y) = \hat{\rho}(x, y)$ pentru orice $x, y \in X$;
   (b) Dacă d este o cvasimetrică invariantă pe $F^a(X, \mathcal{V})$ și $d(x, y) \leq \rho(x, y)$ pentru orice $x, y \in X$, atunci $d(x, y) \leq \hat{\rho}(x, y)$ pentru orice $x, y \in F^a(X, \mathcal{V})$;
   (c) Dacă ρ este o metrică, atunci $\hat{\rho}$ la fel este o metrică;
   (d) Dacă $Y \subseteq X$, $d = \rho|Y$ și $\hat{d}$ este extinderea invariantă maximală a d pe $F^a(Y, \mathcal{V})$, atunci $F^a(Y, \mathcal{V}) \subseteq F^a(X, \mathcal{V})$ și $\hat{d} = \hat{\rho}|F^a(Y, \mathcal{V})$;
   (e) Pentru orice cvasimetrică ρ pe X și orice puncte $a, b \in F^a(X, \mathcal{V})$ există $n \in N$ și reprezentările $a = a_1 a_2 \ldots a_n$, $b = b_1 b_2 \ldots b_n$, astfel încât $a_1, b_1, a_2, b_2, \ldots, a_n, b_n \in X$ și $\hat{\rho}(a, b) = \sum\{\rho(a_i, b_i) : i \leq n\}$. [62]

2. Metoda de extindere a cvasimetricilor pe monoizi liberi în cvasivarietatea completă non-Burnside a monoizilor topologici permite: a construi topologii distincte admisibile a $F^a(X, \mathcal{V})$ pentru orice T_0-spațiu X, a demonstra că monoidul liber topologic $F^a(X, \mathcal{V})$ există pentru orice spațiu X, a stabili că monoidul liber topologic $F(X, \mathcal{V})$ este abstract liber, adică este canonic izomorf cu monoidul abstract liber $F^a(X, \mathcal{V})$ [54, 57, 62].

   Acest fapt rezolvă problemele prezentate de A. I. Maltsev pentru algebrele topologice universale libere [33]. Rezultate similare s-au obținut pentru cvasivarietățile monoizilor semitopologici [62].

3. A fost demonstrat că dacă $\mathcal{V}$ este o cvasivarietate completă non-Burnside de monoizi topologici, atunci X este un spațiu Alexandroff dacă și numai dacă $F(X, \mathcal{V})$ este un spațiu Alexandroff, și X este un spațiu digital dacă și numai dacă $F(X, \mathcal{V})$ este un spațiu digital [61].

   Menționăm că concluziile 1, 2 și 3 nu sunt valabile pentru cvasivarietăți complete Burnside.

4. Pe baza metodelor de extindere a distanței, noțiunile de descompunere paralelă și măsura similarității au fost introduse în spațiul stringurilor [63]. Teorema 3.3.1 descrie relațiile dintre măsura similarității, penalizarea și optimalitatea descompunerilor paralele [56, 58, 59].
5. Diferite relații interesante între distanțele Hamming, Levenshtein și Graev au fost stabilite pe $L(A)$ [46, 47, 48, 55].
6. S-a demonstrat că în clasa tuturor descompunerilor optimale a două stringuri date, măsura maximală a similarității proprii se realizează pe baza descompunerii paralele optimale cu penalitatea minimală (măsura de similaritate minimală); măsura minimală de similaritate proprie se realizează pe descompunerea optimală paralelă cu penalitatea maximală (măsura maximală de similaritate) [51, 63].
7. Au fost propuși algoritmi pentru construirea elementelor mulțimilor de medii ponderate $M_{d_G}(a, b)$ și bisectoare $B_{d_G}(a, b)$ a perechei date de stringuri a și b [49, 64]. A fost ilustrat modul de utilizare a descompunerilor paralele optimale pentru a genera elementele mulțimilor $M_{d_G}(a, b)$, $B_{d_G}(a, b)$, și mulțimii punctelor de mijloc dintre a și b [50, 53, 64].
8. S-a demonstrat că orice medie ponderată a unei perechi de stringuri este generată de careva dintre descompunerile lor paralele optimale [64]. S-a demonstrat, de asemenea, că mulțimea $M_{d_G}(a, b)$ nu este convexă [52].
9. Au fost elaboraţi algoritmi pentru prelucrarea digitală a imaginilor utilizând proprietățile topologiilor digitale dispersate și s-a stabilit că topologia Khalimsky este topologia digitală minimă din clasa tuturor topologiilor simetrice de pe linia discretă $\mathbb{Z}$ [60, 61, 65].

Avantajele și valoarea elaborărilor propuse. Elaborările propuse au o valoare științifică semnificativă datorită gradului lor înalt de originalitate. Rezultatele științifice din această teză au o valoare teoretică și aplicativă în domeniul algebrei, topologiei is cel al informaticii teoretice. De exemplu, metodele de extindere a pseudo-cvasimetricilor pot fi utilizate pentru construirea topologiilor speciale pe monoizi liberi. Metodele de descompunere paralelă, măsura de similaritate, eficiența și penalizarea pot fi aplicate în probleme de analiză a textelor.

Recomandări. Rezultatele obținute pot fi utilizate în diverse domenii și pot avea aplicații practice în algebră și teoria informației. Pe baza concluziilor de mai sus, se recomandă următoarele:

- există un interes special în investigarea cvasimetricilor pe spațiul monoizilor liberi, ca extinderi de cvasimetrici cu proprietăți particulare pe un alfabet. De exemplu, așa cum s-a demonstrat, cvasimetricile sunt strict invariante pe cvasivarietăți rigide. Acest lucru este obișnuit pentru grupuri, dar este foarte rar pentru semigrupuri și monoizi;
- cercetarea rezultatelor poate fi continuată atât din punct de vedere algebric, cât și în scop aplicativ. Cercetarea metricilor pe monoizi este de interes major;
- rezultatele obținute referitoare la descompuneri paralele optimale pot fi utilizate în domeniul alinierii secvențelor de caractere;

- noul algoritm propus pentru construirea mediilor ponderate poate fi mai eficient deoarece ia în considerare simbolul vid, și generează mai multe elemente ale mulțimei M_{d_G} decât algoritmii clasici. Acest fapt, la rândul său, poate fi util în contextul comunicării informațiilor prin canal cu zgomot sau programelor de editare/corectare a textului, în cazul în care are loc pierderea de informație;
- algoritmii de generare a mediilor ponderate și bisectoarei de șiruri de caractere pot fi aplicați in domeniul algoritmilor de analiza a datelor și clusterizare. De exemplu, centrul geometric al unei mulțimi de elemente poate fi calculat ca intersecția elementelor bisectoarelor;
- cercetările ulterioare pot fi continuate prin studiul algoritmilor și proprietăților descompunerilor paralele optimale a trei și mai multor șiruri de caractere;
- conținutul tezei poate servi drept platformă pentru cursurile opționale universitare.

BIBLIOGRAFIE

[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.

PUBLICAȚIILE AUTORULUI LA TEMA TEZEI

[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 reprezentării 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.

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

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.