УДК 519.7
DOI: 10.37069/1683-4720-2022-36-09

©2022. **О.С. Сенченко, М.І. Притула, О.А. Середа**

ПРЕДСТАВЛЕННЯ ДЕТЕРМІНОВАНИХ ГРАФІВ ВИЗНАЧАЛЬНОЮ ПАРОЮ СЛІВ

У роботі запропоновано представлення детермінованих графів множинами слів у алфавіті міток їх вершин. В наш час графи є концептуальним інструментом аналітики, розробкита тестування різноманітного програмного забезпечення. Серед усіх графів виділяють підклас розмічених графів, елементи яких мають мітки з попередньо визначеного алфавіту. Такі графи активно застосовують для описання та моделювання обчислювальних процесів у програмуванні, робототехніці, веріфікації та валідації моделей тощо. Розмічені графи є інформаційним середовищем для мобільних агентів, переміщення яких по графу можна відобразити послідовностями міток вершин – слів в алфавіті міток. Детермінованими називаються графи з розміченими вершинами, у яких в околі кожної вершини всі інші вершини мають попарно різні мітки. Для таких графів у випадку, коли відомі його карта графа (тобто множини вершин і ребер та функція розмітки) та ініціальна вершина, з якої агенти починають свої переміщення, існує однозначна відповідність між послідовністю міток вершин, які відвідує агент, і траєкторією переміщень цього агента на графі. У випадку, коли зовнішньому спостерігачу невідома карта досліджуваного детермінованого графа, переміщення агентів можуть бути організовані в такий спосіб, щоб на основі їх аналізу спостерігач одержав шукану інформацію щодо структури графа (наприклад, карту графа, найкоротші шляхи між вершинами, порівняння досліджуваного графа з графом-еталоном). Такий аналіз може суттєво спростити лінгвістичне представлення детермінованого графа – співставлення графу однієї чи декількох скінченних множин слів у алфавіті міток вершин графа, за якими можна відновити граф. У даній роботі запропоновано представлення детермінованих графів визначальною парою множин слів, перша компонента якої описує цикли графа, а друга – його висячі вершини. Це представлення є аналогом системи визначальних співвідношень для повністю визначених автоматів. Наведено алгоритм, який за довільною парою множин або будує детермінований граф, для якого ця пара є визначальною, або сповіщує, що це зробити неможливо. Також наведено алгоритм побудови канонічної визначальної пари для детермінованого графа та знайдено чисельні оцінки цієї пари для детермінованих графів із відомою кількістю вершин та ребер. Окреслено подальші напрямки дослідження за цією тематикою. Результати дозволять використовувати нові методи та алгоритми для розв'язання задач аналізу графів з розміченими вершинами.

MSC: 68R10.

Ключові слова: *детерміновані графи, представлення, визначальна пара.*

1. Вступ.

В наш час графи є концептуальним інструментом аналітики, розробки та тестування різноманітного програмного забезпечення. Серед усіх графів виділяють підклас розмічених графів, тобто графів, у яких вершини та/або ребра мають деякі мітки з попередньо визначеного алфавіту. За допомогою таких графів природньо моделюються різноманітні процеси в сучасних інформаційних системах. При цьому розмічені графи, у яких мітки знаходяться на ребрах (LTS, зважені графи, скінченні автомати тощо) досліджені в значно більшій мірі, ніж графи з розміченими вершинами. Але існує багато обчислювальних процесів у програмуванні [1], робототехніці [2], верифікації та валідації моделей [3], які зручно відображати саме графами з розміченими вершинами.

У роботі [4] запропоновано представлення скінченних, всюди визначених автоматів без виходів системами визначальних співвідношень – множинами пар слів спеціального вигляду, що задають еквівалентні траєкторії в автоматі. За допомогою такого представлення було розв'язано задачу характеризації – визначення, чи є дана множина пар слів системою визначальних співвідношень для заданого автомата без безпосередньої побудови самого автомата. Природним чином постає задача розповсюдження такого представлення на інші об'єкти дискретної математики, зокрема на графи з розміченими вершинами, принаймні на деякий підклас таких графів. На наш погляд, найбільш зручними для такого розповсюдження є детерміновані графи – клас графів з розміченими вершинами, у яких в околі кожної вершини всі вершини мають попарно різні мітки. Вперше такі графи було формалізовано в роботі [5]. Доцільність такого розповсюдження випливає з природнього взаємозв'язку між детермінованими графами та мовами в алфавіті міток вершин (для детермінованих графів траєкторія руху однозначно відновлюється за послідовністю міток вершин). Цей взаємозв'язок дозволяє застосувати методи теорії автоматів до аналізу розмічених графів.

В даній роботі узагальнено всі одержані нами результати за цією темою та окреслено напрямки подальших досліджень, пов'язаних з представленням детермінованих графів визначальною парою слів.

2. Основні означення.

У роботі розглядаються неорієнтовані, скінченні, непорожні, зв'язні, звичайні (такі, що не містять петлі та кратні ребра) графи з розміченими вершинами $G = (V, E, X, \xi(V))$, де V – множина вершин графа, E – множина його ребер, $\xi(V) : V \to X$ – всюди визначена функція розмітки вершин графа символами скінченного алфавіту $X = \{x_1, \ldots, x_p\}$. Далі такі графи будемо скорочено називати терміном «розмічений граф». Значення функції розмітки для вершини v будемо називати її міткою. Множину всіх слів скінченної довжини в алфавіті X (у тому числі порожнє слово ε) будемо позначати через X^*. Нехай $p = x'_1 \ldots x'_k$ ($x'_i \in X$), тоді довжину слова p будемо позначати через $d(p)$. Через $E(v)$ позначатимемо множину вершин, що є суміжними вершині v: $v' \in E(v) \leftrightarrow (v, v') \in E$.

У залежності від додаткових умов, що накладаються на функцію розмітки вершин, з множини всіх розмічених графів виділяють окремі підкласи. У цій роботі розглядаються так звані детерміновані графи, у яких для будь-яких вершин $v, v', v'' \in V$ з $(v, v'), (v, v'') \in E$ та $\xi(v') = \xi(v'')$ випливає рівність $v' = v''$ [5]. Змістовно кажучи, наведене означає, що у детермінованих графів будь-які дві вершини, що одночасно суміжні деякій вершині, мають різні мітки. Зауважимо, що на відміну від так званих сильнодетермінованих графів [5], у детермінованих графів можуть існувати суміжні вершини з однаковими мітками.

Зафіксуємо деяку вершину $v_0 \in V$ Д-графа $G = (V, E, X, \xi(V))$, яку далі будемо називати ініціальною. У цій роботі ми будемо розглядати саме детерміновані графи з ініціальною вершиною та коротко їх називати «Д-графами», цю вершину за необхідності ми будемо виділяти у позначенні Д-графа: $G = (V, E, X, \xi(V), v_0)$.

Ізоморфізмом Д-графів $G_1 = (V_1, E_1, X, \xi_1(V_1), v')$ і $G_2 = (V_2, E_2, X, \xi_2(V_2), v'')$ з однаковою множиною міток вершин X називаємо взаємно однозначну відповідність $\varphi : V_1 \leftrightarrow V_2$ між множинами їх вершин, що зберігає їх ініціальні вершини, відношення суміжності вершин та функцію розмітки вершин: $\varphi(v') = v''$, $\forall v', v'' \in V_1 \quad ((v', v'') \in E_1 \leftrightarrow (\varphi(v'), \varphi(v'')) \in E_2)$ та $\forall v \in V_1 \quad \xi_1(v) = \xi_2(\varphi(v))$; ізоморфізм графів G_1 та G_2 будемо позначати $G_1 \cong G_2$ або $G_1 \cong_\varphi G_2$. Зафіксуємо вершину v'_1 Д-графа $G = (V, E, X, \xi(V))$. Оскільки шляху $p = v'_1 \ldots v'_k$ однозначно відповідає слово в алфавіті міток $\xi(p) = \xi(v'_1)\,\xi(v'_2) \ldots \xi(v'_k)$ [5], то надалі шлях p розглядатимемо як $\xi(p)$ та позначатимо рівністю $v'_1 p = v'_k$.

Шлях $p = x'_1 x'_2 \ldots x'_k$ назвемо припустимим для вершини $v'_1 \in V$, якщо $\xi(v'_1) = x'_1$, та існують такі вершини $v'_2, \ldots, v'_k \in V$, що $\xi(v'_2) = x'_2, \ldots, \xi(v'_k) = x'_k$, і $(v'_1, v'_2), \ldots, (v'_{k-1}, v'_k) \in E$. Слово $x'_k \ldots x'_1$ будемо називати реверсом слова $p = x'_1 \ldots x'_k$ та позначати його через p^{-1}, слова p та p^{-1} назвемо взаємнозворотними. Слово p, для якого виконується рівність $vp = v$, назвемо циклічним для вершини v. Слово $q = x''_1 \ldots x''_m$ назвемо початковим відрізком слова $p = x'_1 \ldots x'_k$ у випадку, якщо $m \leq k$ та виконуються рівності $x''_1 = x'_1, \ldots, x''_m = x'_m$, цей факт позначатимемо через $q \subseteq p$.

Всі вершини заданого Д-графа $G = (V, E, X, \xi(V), v_0)$ можна розбити на два класи за такою процедурою: видалимо з G усі висячі вершини, відмінні від ініціальної, разом із ребрами, що є інцидентними цим вершинам. Цю процедуру будемо повторювати до тих пір, поки це можливо. Одержаний у такий спосіб граф $B(G)$ назвемо базою графа G, вершини, що входять до $B(G)$, назвемо базовими, а ті вершини G, що не входять до $B(G)$, – вільними. Неважко бачити, що база графа будується однозначно за скінченну кількість кроків, база $B(G)$ може складатися з однієї ініціальної вершини (у випадку, коли G є деревом), а у випадку, коли $B(G)$ містить хоча б одну вершину, яка відмінна від ініціальної, то вона повинна містити хоча б один простий цикл. Звідси випливає, що кожна базова вершина, яка є відмінною від ініціальної вершини v_0, суміжна щонайменше двом базовим вершинам. Вільну вершину v назвемо прямим предком вільної вершини v', якщо для будь-якого шляху p', такого, що $v_0 p' = v'$, існує такий початковий відрізок $p \subseteq p'$, що $v_0 p = v$, цей факт позначатимемо через $v \preceq v'$. Звернемо увагу, що для будь-якої вільної вершини v виконується $v \preceq v$.

3. Лінгвістичне представлення Д-графів.

Для дослідження структури розмічених графів часто використовують один або декілька мобільних агентів з обмеженою пам'яттю, які поміщають на досліджуваний граф [6]. Ці агенти можуть сповіщати спостерігачу та/або іншим агентам певну інформацію щодо локальних околів вершин, на яких вони наразі знаходяться. На основі цієї інформації та/або інструкцій спостерігача агенти можуть переміщатися по ребрах графа. Переміщення агентів визначають послідовності міток вершин, які вони відвідали. Зважаючи на це, Д-графи є дуже зручними для такого дослідження, оскільки, якщо спостерігач має карту графа (множини V, E, X та функцію $\xi(V)$) та знає, на яку вершину досліджуваного Д-графа був поміщений агент на початку дослідження (тобто, ініціальну вершину), то за цією послідовністю міток може бути однозначно відновлена траєкторія переміщення агента по графу. У випадку, коли спостерігачу не відома карта досліджуваного графа, переміщення агентів можуть бути організовані в такий спосіб, щоб на основі їх аналізу спостерігач одержав шукану інформацію про структуру графа (наприклад, складення карти графа, пошуку у ньому найкоротших шляхів, порівняння досліджуваного графа з графом-еталоном).

Такий аналіз може суттєво спростити лінгвістичне представлення Д-графа, під яким ми розуміємо співставленню графу однієї чи декілька скінченних множин слів у алфавіті міток вершин графа, що описують можливі (та/або неможливі) траєкторії руху в графі. Для такого представлення треба визначитися з кількістю таких множин слів (бажано, щоб така кількість була якомога меншою) та роллю, яку відіграє кожна з цих множин слів у графі (які структурні елементи графа описує кожна множина). При цьому, ці множини треба обрати у такий спосіб, щоб можна було винайти універсальний алгоритм, який кожному набору цих множин або однозначно співставляє детермінований граф, або сповіщає, що це зробити неможливо. Також треба винайти алгоритм переходу від будь-якого Д-графа до сказаного вище набору множин.

Далі вважаємо, що всі шляхи в графі починаються з ініціальної вершини. Будемо робити представлення Д-графа $G = (V, E, X, \xi(V), v_0)$, у якого $\xi(v_0) = (x')$ парою $\{C, L\}(x')$ скінченних множин слів $C, L \in X^*$, для якої всі слова з C та L є припустимими для v_0, слова множини C описують цикли графа $G = (V, E, X, \xi(V), v_0)$, а слова L описують його висячі вершини. Далі в роботі під терміном «пара» $\{C, L\}(x')$ розуміємо саме такі дві скінченні множини C та L, що всі слова множини C починаються та закінчуються символом x', а всі слова множини L починаються з цього символа. При цьому множини C та L будемо називати компонентами пари $\{C, L\}(x')$.

Нехай $G = (V, E, X, \xi(V), v_0)$ – деякий Д-граф. Зафіксуємо на X лінійний порядок $<$: $x_1 < x_2 < \cdots < x_p$. Визначимо на X^* лінійний порядок $\preceq$:

1) для будь-якого $x_i \in X$, де $1 \le i \le p$, покладемо $x_i \preceq x_i$;

2) для будь-яких $p, q \in X^*$, якщо $d(p) < d(q)$, то $p \preceq q$;

3) для будь-яких $p, q \in X^*$, якщо $p = x'_1 \dots x'_s$, $q = x''_1 \dots x''_s$, $x'_1 = x''_1$, …, $x'_{k-1} = x''_{k-1}$ та $x'_k < x''_k$, то $p \preceq q$.

Неважко бачити, що порядок $\preceq$ на множині X^* визначається лінійним порядком $<$ на множині X.

Нехай $G = (V, E, X, \xi(V), v_0)$ – довільний граф з розміченими вершинами. Редукцією G назвемо процедуру, яка перетворює G у Д-граф $[G]$ за таким алгоритмом (АР):

0) Покладемо $V' = V, E' = E, \xi'(V') = \xi(V)$, позначимо $G' = (V', E', X, \xi'(V'), v_0)$.

1) Для кожної вершини $v \in V'$ знаходимо найменше за $\preceq$ слово w, що відповідає шляху від v_0 до v. Пов'язуємо v з w.

2) Впорядковуємо вершини графа G' за пов'язаними з ним словами за порядком

$\preceq$. Поточною вершиною v_c призначимо першу за вказаним порядком вершину, а поточною міткою призначимо $x = x_1$.

3) Якщо існують такі різні вершини $v'_1, \ldots, v'_q$, для яких одночасно виконуються умови $v'_1, \ldots, v'_q \in E'(v_c)$ та $\xi'(v'_1) = \ldots = \xi'(v'_q) = x$, то позначимо $U = \{v'_1, \ldots, v'_q\}$ та виконуємо таку послідовність дій:

3.1. до множини V' додаємо нову вершину v' та покладемо $\xi'(v') = x$;

3.2. з множини E' вилучаємо ребра (v_i, v_j), де $v_i, v_j \in U$;

3.3 для кожного ребра (v_i, v_j), де $v_i \in U$, до множини E' додаємо ребро (v', v_j);

3.4. з множини E' вилучаємо ребра (v_i, v_j), де $v_i \in U$, з множини V' вилучаємо вершини v, де $v \in U$;

3.5 якщо $v_0 \in U$, то вершину v' перейменовуємо на v_0 та вважаємо далі цю вершину ініціальною;

3.6. вилучаємо повтори ребер, залишаючи по одному екземпляру;

3.7. переходимо на крок 1.

4) Якщо існує $x' \in X$, який є наступним за x за порядком $<$ (тобто, $x \neq x_p$), то покладемо $x = x'$ та переходимо до кроку 3.

5) Якщо існує вершина v, що є наступною для v_c, то покладемо $v_c = v$, $x = x_1$ та переходимо на крок 3, у протилежному випадку алгоритм АР завершує свою роботу та $[G] = G'$.

Неважко бачити, що виконання алгоритму АР закінчується за скінченну кількість кроків, його результат визначається однозначно та цей результат є Д-графом.

Зауваження. При виконанні цього алгоритму існує певна неоднозначність з іменами нових вершин (окрім випадку, за яким новостворену вершину v' було перейменовано на v_0 на кроці 3.5), проте, для нас є несуттєвими імена усіх інших вершин (окрім ініціальної) у графі. За цієї обставини, якщо було виконано крок 3 алгоритму АР, ми не можемо говорити про рівність графа $[G]$ якомусь іншому графу, а можемо говорити про ізоморфність $[G]$. Також неважко бачити, що $[G] = G$ тоді та лише тоді, коли G є Д-графом.

Далі під терміном «пара» $\{C, L\}(x')$ розуміємо дві множини C та L, елементи яких – слова в алфавіті X – задовольняють таким умовам:

а) кожне слово множини C розпочинається та закінчується символом x';

б) довжина кожного слова множини C більша 2;

в) кожне слово множини L розпочинається символом x'.

Визначимо алгоритм (АП), який за заданою парою $\{C, L\}(x')$ або будує Д-граф $G(\{C, L\}(x'))$, або показує, що за цією парою Д-граф побудувати неможливо.

0) Покладемо, що спочатку граф $G(\{C, L\}(x'))$ складається з однієї вершини v_0 з міткою $\xi(v_0) = x'$.

1) У відповідність кожному слову $p^i = x' x^i_1 \ldots x^i_n x' \in C$ додаємо до графа вершини $v^i_1, \ldots, v^i_n$ з мітками відповідно $x^i_1, \ldots, x^i_n$ та ребра $(v_0, v^i_1), (v^i_1, v^i_2), \ldots, (v^i_{n-1}, v^i_n)$, (v^i_n, v_0). Після кожного такого додавання робимо редукцію одержаного графа.

2) У графі, який ми одержали після виконання кроку 1, можуть існувати висячі вершини. Кожну таку вершину v, що є відмінною від ініціальної вершини v_0,

поміщаємо для подальшого аналізу у множину Q, яка на початку була порожня.

3) Для кожного слова $p^j = x'x_1^j \dots x_n^j \in L$ виконуємо таку послідовність дій: додаємо вершини $v_1^j, \dots, v_n^j$ з мітками відповідно $x_1^j, \dots, x_n^j$, ребра $(v_0, v_1^j), \dots, (v_{n-1}^j, v_n^j)$ та робимо редукцію одержаного графа. Якщо в результаті редукції вершина $v_0 p_j$ не є висячею, то вважаємо, що граф $G(\{C, L\}(x'))$ не існує, і процедура завершується. Якщо ж в результаті редукції з'явилась висяча вершина v, відмінна від вершини $v_0 p_j$, то цю вершину поміщаємо для подальшого аналізу у множину Q. Після цього переходимо до наступного слова компоненти L.

4) Для кожної вершини $v \in Q$ розглядаємо всі слова компоненти L: якщо не існує такого $p \in L$, що вершина v є прямим предком вершини $v_0 p$, то вважаємо, що граф $G(\{C, L\}(x'))$ не існує.

5) Розглядаємо всі слова множини L: якщо існує таке $p \in L$, що вершина $v_0 p$ не є висячою, то вважаємо, що граф $G(\{C, L\}(x'))$ не існує.

Якщо в результаті виконання цієї процедури за парою $\{C, L\}(x')$ можливо побудувати граф $G(\{C, L\}(x'))$, то таку пару назвемо правильною. Правильну пару $\{C, L\}(x')$ назвемо *визначальною* для Д-графа G, якщо $G(\{C, L\}(x')) \cong G$. Зауважемо, що таке завдання Д-графа визначальною парою в деякій мірі (етапами 1 та 3) є аналогічним представленню автоматів системою визначальних співвідношень, запропонованому у [4].

Доцільність введення другого, четвертого та п'ятого кроків до алгоритму АП обумовлено намаганням чітко розділити функції компонентів пари: передбачається, що слова компоненти C описують базові вершини графа, а слова компоненти L – його вільні вершини, причому кожне слово $p \in L$ описує висячу вершину $v_0 p$, а для кожної вільної висячої вершини v потрібне існувати хоча б одне таке слово $p \in L$, що $v_0 p = v$ (зауважимо, що у випадку, коли вершина v_0 є висячею, то її не потрібно описувати у компоненті L). У випадку, коли слова множини C задають деякі вільні вершини графа (тобто після виконання кроку 2 множина Q не є порожньою), ці вільні вершини обов'язково повинні бути описаними словами компоненти L (див. крок 4), а якщо такого опису нема, то вважаємо, що за парою $\{C, L\}(x')$ граф $G(\{C, L\}(x'))$ побудувати неможливо.

Розглянемо дію окремих етапів наведеної вище процедури на такому прикладі.

Приклад 1. Нехай $C = \{153521, 152431\}$, $L = \{1342531, 123241, 13412, 1523\}(1)$. На рис. 1 а зображено граф, який одержуємо після виконання етапів 1 та 2. Вершина v', що виділена, є висячею та відмінною від ініціальної вершини v_0, тому цю вершину поміщаємо до множини Q.

На рис. 1 б зображено граф, який одержуємо після виконання етапу 3. Після розгляду слова 123241 вершина v'', що виділена, є висячею та відмінною від вершини v_0, тому цю вершину поміщаємо до множини Q.

Далі робимо аналіз вершин v' та v'', що входять до множини Q. Оскільки вершина v' є прямим предком вершини u', $1342531 \in L$ та $v_0 1342531 = u'$, а також вершина v'' є прямим предком самої себе, $1523 \in L$ та $v_0 1523 = v''$ (рис. 1 в), то переходимо до виконання етапу 5.

Оскільки вершина v (рис. 1 г) не є висячою, $v_0 123241 = v$ та $123241 \in L$, то за умовою етапу 5 граф $G(\{C, L\}(1))$ не існує.

Зауважимо, що для пари $\{C, L'\}(1)$, де $L' = (L \setminus \{123241\}) \bigcup \{1232412\}$ граф $G(\{C, L'\}(1))$ існує, він зображений на рис. 1 г.

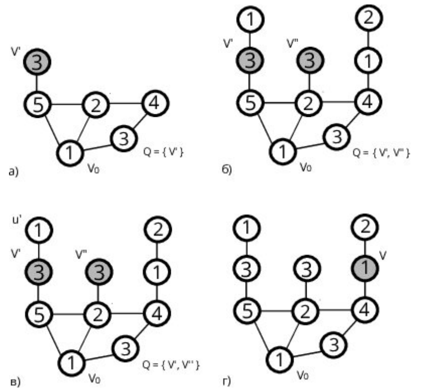

Рис. 1. Ілюстрація процедури побудови Д-графа за заданою парою

4. Канонічна визначальна пара.

Нехай $G = (V, E, X, \xi(V), v_0)$ – деякий Д-граф, та $V = \{v_0, \ldots, v_{n-1}\}$. Визначимо допоміжну множину слів в алфавіті X. Базисом досяжності $\mathcal{V}_G$ назвемо таку множину слів $\{w_1, w_2, \ldots, w_n\}$, де для кожної вершини v_i існує таке слово $w_i \in \mathcal{V}_G$, що виконується $v_0 w_i = v_i$, та для будь-якого $w \neq w_i$ з $v_0 w = v_i$, випливає $w_i \preceq w$. Кістякове дерево графа G, яке визначається базисом $\mathcal{V}_G$, позначимо $T(\mathcal{V}_G)$ або $T(G, v_0)$. Для кожної вершини $v \in V$ через sp_v позначимо таке слово з $\mathcal{V}_G$, що $v_0 sp_v = v$.

Опишемо алгоритм (АК) побудови для Д-графа G визначальної пари $\{\Sigma_G, \Lambda_G\}$, яку, завдяки деяким її властивостям, ми назвали канонічною.

Спочатку покладемо $\Sigma_G = \emptyset$ і $\Lambda_G = \emptyset$. Якщо граф G складається з однієї вершини v_0, то покладемо $\Sigma_G = \emptyset$, $\Lambda_G = \emptyset$ і $\mathcal{V}_G = \{\xi(v_0)\}$.

Нехай граф G містить більше ніж одну вершину. Спочатку до множини Λ_G додаємо усі слова $w \in \mathcal{V}_G$ такі, що вершина $v_0 w$ є висячою вершиною графа G. Після цього для кожної двійки слів $p, q \in \mathcal{V}_G \setminus \Lambda_G$, якщо жодне з них не є початковим відрізком іншого і $v_0 p q^{-1} = v_0$, то додаємо до множини Σ_G одне з двох слів pq^{-1} або qp^{-1}, яке є меншим за порядком $\preceq$ (далі вершини $v_0 p$ та $v_0 q^{-1}$ називатимемо твірними для того слова pq^{-1} або qp^{-1}, що було додане до Σ_G). Після цього прибираємо усі повтори слів у Σ_G, залишаючи по одному екземпляру.

Опишемо деякі властивості канонічної визначальної пари, що безпосередньо випливають з алгоритму АП.

Теорема 1. *Нехай* $\{\Sigma_G, \Lambda_G\}$ *– канонічна визначальна пара Д-графа* $G = (V, E, X, \xi(V), v_0)$.

1) Якщо хоча б одна компонента $\{\Sigma_G, \Lambda_G\}$ *не є порожньою множиною, то* *для кожної вершини $v \in V$існують такі $z_1, z_2, z_3 \in X^*$, що виконується хоча б одне з тверджень: а) $sp_v z_1 \in \Lambda_G$; б) $sp_v z_2 \in \Sigma_G$; в) $z_3(sp_v)^{-1} \in \Sigma_G$.*

2) Нехай (v_1, v_2) – деяке ребро графа G, $\xi(v_1) = x'_1$ та $\xi(v_2) = x'_2$. Тоді існують такі $z_1, z_2 \in X^*$, що виконується хоча б одне з тверджень: а) $sp_{v_1} x'_2 z_1 \in \Sigma_G \bigcup \Lambda_G$; б) $sp_{v_2} x'_1 z_2 \in \Sigma_G \bigcup \Lambda_G$.

3) Якщо $\Sigma_G \neq \emptyset$, то для кожного $\sigma \in \Sigma_G$ існують такі $p, q \in X^*$, що $pq = \sigma$, та $p \in \mathcal{V}_G$ і $q^{-1} \in \mathcal{V}_G$.

4) Нехай (v_1, v_2) – деяке ребро графа G, яке не належить кістяковому дереву $T(G, v_0)$. Тоді існує єдине слово $\sigma \in \Sigma_G$, що ребро (v_1, v_2) входить до шляху $v_0\sigma$.

5) Якщо $\Sigma_G \neq \emptyset$, то для кожного $\sigma \in \Sigma_G$ виконується нерівність $d(\sigma) \geq 4$.

6) Якщо $\Sigma_G \neq \emptyset$, то для кожного $\sigma \in \Sigma_G$ знайдуться такі $p = p'x_1, q \in X^*(x_1 \in X)$, $x_2, x_3 \in X$, $x_2 \neq x_3$, що $\sigma = px_2qx_3p^{-1}$ та $v_0p = v_0px_2qx_3x_1$.

7) Якщо v_1, v_2 – твірні вершини деякого слова $\sigma \in \Sigma_G$, то $(v_1, v_2) \notin T(G, v_0)$ та $\mid d(sp_{v_1}) - d(sp_{v_2}) \mid \leq 1$.

Змістовно кажучи, перше та друге твердження стверджують, що пара $\{\Sigma_G, \Lambda_G\}$ містить певну інформацію про кожну вершину та ребро графа G, третє твердження встановлює своєрідну мінімальність початкових та кінцевих відрізків слова $\sigma \in \Sigma_G$, четверте твердження вказує, що кожне ребро, яке не належить кістяковому дереву $T(\mathcal{V}_G)$, описується окремим словом з Σ_G, п'яте твердження показує, що σ містить простий цикл; який, як встановлює шосте твердження, не може бути представлений у вигляді pxp^{-1} для будь-яких $p \in X^*$, $x \in X$.

З алгоритмів АП та АК можна бачити, що пара $\{\Sigma_G, \Lambda_G\}$ є правильною (для кожного слова $\lambda \in \Lambda_G$ вершина $v_0\lambda$ є висячою, а множина Q є порожньою). Також справедлива

Теорема 2. $G(\{\Sigma_G, \Lambda_G\}) \cong G$.

Для доведення цієї теореми була досліджена покрокова побудова графа $G(\{\Sigma_G, \Lambda_G\})$ за словами множин Σ_G та Λ_G. Для цього покладемо $\Sigma_G = \{\sigma_1, \ldots, \sigma_k\}$ та $\Lambda_G = \{\lambda_1, \ldots, \lambda_q\}$. Введемо сімейство множин M_i $(0 \leq i \leq k+q)$ у такий спосіб: $M_0 = \varnothing$, $M_1 = \{\sigma_1\}$, $M_2 = \{\sigma_1, \sigma_2\}$, ..., $M_k = \{\sigma_1, \ldots, \sigma_k\} = \Sigma_G$, $M_{k+1} = \{\sigma_1, \ldots, \sigma_k, \lambda_1\}$, ..., $M_{k+q} = \{\sigma_1, \ldots, \sigma_k, \lambda_1, \ldots, \lambda_q\} = \Sigma_G \bigcup \Lambda_G$. За словами цього сімейства множин вводяться два сімейства графів: G_i та H_i. Графи сімейства G_i є підграфами графа G, що визначаються вершинами та ребрами слів з відповідної множини M_i, а графи сімейства H_i побудовані за словами з відповідної множини M_i за допомогою алгоритму АП. За цим алгоритмом $H_{k+q} = G(\{\Sigma_G, \Lambda_G\})$, а з пршого та другого тверджень теореми 1 випливає рівність $G_{k+q} = G$. Індукцією за i $(0 \leq i \leq k+q)$ показано, що $H_i \cong G_i$. Доведення є досить великим за обсягом, тому в цій роботі ми його не наводимо.

5. Метричні властивості компонент канонічної визначальної пари.

Наведемо всі знайдені нами чисельні оцінки компонент канонічної визначальної пари. Нехай далі Д-граф $G = (V, E, X, \xi(V), v_0)$ містить n вершин та m ребер $(n - 1 \leq m \leq \frac{n \cdot (n-1)}{2})$. Далі позначення $\lceil x \rceil$ означає найменьше натуральне число, що не меньше за x. У [7] було знайдено та доведено точне значення потужності першої компоненти та досяжні оцінки потужності другої компоненти канонічної визначальної пари.

Теорема 3.

1) $|\Sigma_G| = m - n + 1$.

2) Якщо $m = n - 1$ *(тобто* G *є деревом), то* $1 \leq |\Lambda_G| \leq n - 1$. *Якщо ж* $m > n - 1$, *то* $0 \leq |\Lambda_G| \leq n - \lceil \frac{3}{2} + \sqrt{\frac{9}{4} - 2 \cdot n + 2 \cdot m} \rceil$, *причому всі оцінки є досяжними.*

Також було знайдено досяжні оцінки $\|\Sigma_G\|$ значення сумарної довжини усіх слів (об'єм) першої компоненти визначальної пари Д-графа G. Якщо G є деревом, то $\Sigma_G = \emptyset$, тому у цьому випадку $\|\Sigma_G\| = 0$. Для випадку, коли G не є деревом, справедлива

Теорема 4. *Якщо* $m > n - 1$, *то*
$4(m - n + 1) \leq \|\Sigma_G\| \leq 4(m - n + 1)\left(n - \lceil \frac{3}{2} + \sqrt{\frac{9}{4} - 2 \cdot n + 2 \cdot m} \rceil + 2\right)$, *причому ці оцінки є досяжними.*

Доведення. Нижня оцінка безпосередньо випливає з першого твердження теореми 3 та п'ятого твердження теореми 1. Ця оцінка є досяжною для графа, всі вершини якого суміжні ініціальній вершині. У такого графа Σ_G складається з $m - n + 1$ слів довжини 4.

Доведемо верхню оцінку. За першим твердженням теореми 3, $|\Sigma_G| = m - n + 1$, покладемо $p = m - n + 1$. За третім твердженням теореми 1, кожне слово $\sigma \in \Sigma_G$ може бути представленим парою твірних вершин u, w, тому $d(\sigma) = d(sp_u) + d(sp_w)$. Нехай V' – множина вершин графа G, що є твірними для хоча б одного слова з Σ_G. Таким чином, знаходження верхньої оцінки $\|\Sigma_G\|$ перетворюється у знаходження максимально можливих за довжиною значень шляхів $sp_v (v \in V')$. Зважаючи на сьоме твердження теореми 1 та той факт, що збільшення кількості вершин у V' при фіксованій кількості вершин у G, призводить до зменшення максимально можливої довжини їх $sp_v (v \in V')$, найбільше значення $\|\Sigma_G\|$ можливо у графі, у якому потужність V' є мінімально можливою для утворення p слів у Σ_G, всі вершини з V' суміжні деякій фіксованій вершині v', та іншим вершинам з V' та не суміжні жодній вершині з $V \setminus (V' \bigcup \{v'\})$. У такому графі всі прості цикли є трикутниками, одна вершина якого є v', а дві інші вершини належать множині V', $d(sp_v) = d(sp_{v'}) + 1$ $(v \in V')$ та $\|\Sigma_G\| = 2p(d(sp_{v'} + 1))$; найбільш можливе значення $\|\Sigma_G\|$ буде у графа, у якого значення $y = d(sp_{v'})$ буде максимально можливим.

Для знаходження максимально можливого значення y опишемо структуру графа G' з n вершинами та m ребрами, для якого кількість висячих вершин, що не належать базі $B(G')$, буде максимальною. Нехай у графі G' t вершин належать його базі. Неважко бачити, що максимальна кількість вільних висячих вершин у G' дорівнює $n - t$, це можливо у випадку, при якому всі вільні вершини графа є висячими (як приклад, всі вільні вершини такого графа є суміжними ініціальній вершині), мінімально можливе значення t далі будемо позначати через $\delta(m, n)$. Для знаходження $\delta(m, n)$ використаємо той очевидний факт, що найбільша кількість простих циклів та найбільша потужність першої компоненти канонічної визначальної пари для графа G з n вершинами є у повного графа K_n. Тому $\delta(m,n)$ дорівнює такому мінімальному значенню z, що число p не переважає потужність $|\Sigma_{K_z}|$. Оскільки у K_z кількість ребер дорівнює $\frac{z\cdot(z-1)}{2}$, то $|\Sigma_{K_z}| = \frac{z\cdot(z-1)}{2} - z + 1 = (\frac{z}{2} - 1)\cdot(z-1)$. Іншими словами, значенням $\delta(m,n)$ буде таке мінімальне натуральне число z, що задовільняє нерівності $m-n+1 \leq (\frac{z}{2}-1)\cdot(z-1)$. З урахуванням натуральності m, n та z шукане значення $\delta(m,n) = \lceil\frac{3}{2}+\sqrt{\frac{9}{4} - 2\cdot n + 2\cdot m}\rceil$. Оскільки $\delta(m,n) = |V'|+1$, та існують $n-\delta(m,n)$ вершин, що не належать множині $V' \bigcup \{v'\}$, то максимально можливе значення y буде $n - \delta(m,n) + 1$, тому $\|\Sigma_G\| \leq 4(m-n+1)\left(n - \lceil\frac{3}{2} + \sqrt{\frac{9}{4} - 2\cdot n + 2\cdot m}\rceil + 2\right)$. $\square$

Зауваження. Зауважимо, що наведені формули для верхньої та нижньої оцінок цієї теореми справедливі також для випадку $m = n-1$, за яким вони дорівнюють 0, що збігається зі значенням об'єму першої компоненти канонічної визначальної пари для дерева, яке було знайдено вище.

Знайдено оцінки $\|\Lambda_G\|$ для випадку, коли граф G є деревом.

Теорема 5. *Нехай Д-граф $G = (V, E, X, \xi(V), v_0)$ є деревом та містить n вершин. Тоді $n \leq \|\Lambda_G\| \leq \lceil\frac{n^2+2n}{4}\rceil$, причому ці оцінки є досяжними.*

Доведення. Оскільки G є деревом, то $\Sigma_G = \emptyset$. Тоді, з урахуванням другого твердження теореми 1, для кожного ребра (v_1, v_2) графа G повинно існувати таке слово $\lambda \in \Lambda_G$, що (v_1, v_2) входить до шляху $v_0\lambda$. Оскільки шлях з k ребер складається з $k+1$ символів алфавіта X, то найменш можлива сумарна кількість символів у Λ_G, яка необхідна для входження у слова з Λ_G всіх $n-1$ ребер графа G, становить $n-1+1$, що доводить нижню оцінку.

Доведемо верхню оцінку. Нехай граф G має q висячих вершин, які відмінні від вершини v_0, та, відповідно $|\Lambda_G| = q$ ($q = \overline{1,\ldots,n-1}$). Неважко бачити, що максимально можливе значення $\|\Lambda_G\|$ буде у графа, у якого всі висячі вершини, відмінні від v_0 суміжні деякій фіксованій вершині v', та значення $d(sp_{v'})$ буде максимально можливим, тобто $d(sp_{v'}) = n - q$. Тоді довжина кожного слова з Λ_G дорівнює $n-q+1$, отже, $\|\Lambda_G\| \leq q(n-q+1)$. Максимальне значення $\|\Lambda_G\|$ буде у випадку $q = \frac{n+1}{2}$. З урахуванням натуральності числа q, для непарних n $q = \frac{n+1}{2}$, а для парних n, $q = \frac{n}{2}$, або $q = \frac{n}{2}+1$. Об'єднуючи ці випадки, одержемо $\|\Lambda_G\| \leq \lceil\frac{n^2+2n}{4}\rceil$. $\square$

6. Подальші напрямки дослідження.

Наведене лінгвістичне представлення Д-графа визначальною парою слів може бути корисним при розв'язанні багатьох прикладних задач. При цьому природним чином виникає цілий ряд задач, пов'язаних з цим лінгвістичним представленням, серед яких автори виділяють наступні:

- знаходження взаємозв'язків між графами G та $[G]$;

- розв'язання задачі характеризації пари: для заданого Д-графа та заданої пари визначити, чи є ця пара визначальною для графа без безпосередньої побудови графа за парою;

- знаходження необхідних та достатніх умов для множин C та L, за яких пара $\{C, L\}$ є правильною;

- оптимальний вибір ініціальної вершини, за яким метричні властивості компонент канонічної визначальної пари будуть мінімальними;

- ефективна побудова за $\{\Sigma_G, \Lambda_G\}$ найкоротших за порядком $\preceq$ шляхів між двома довільними вершинами графа G;

- застосування наведених у цій роботі понять та алгоритмів для сильнодетермінованих [5] графів.

7. Висновки.

У роботі запропоновано лінгвістичне представлення детермінованого графа визначальною парою слів. Наведено алгоритм, який за довільною парою множин або будує Д-граф, для якого ця пара є визначальною, або сповіщує, що це зробити неможливо. Також наведено алгоритм побудови канонічної визначальної пари для Д-графа. Знайдено чисельні оцінки цієї пари для Д-графів із відомою кількістю вершин та ребер – потужність першої компоненти пари, мінімальні й максимальні досяжні оцінки потужності другої компоненти пари та об'єму її першої компоненти. Також знайдено мінімальні й максимальні досяжні оцінки об'єму другої компоненти канонічної визначальної пари для випадку, коли граф є деревом. Окреслено подальші напрямки дослідження за цією тематикою. Результати дозволять використовувати нові методи та алгоритми для розв'язання задач анализу графів з розміченими вершинами.

Цитована література

1. *Okhotin A.* Graph-Walking Automata: From Whence They Come, and Whither They are Bound. – In: Hospodár M., Jirásková G. (eds) Implementation and Application of Automata. CIAA 2019. – Lecture Notes in Computer Science, 2019, vol 11601. – Springer, Cham.
2. *Dudek G., Jenkin M.* Computational Principles of Mobile Robotics. – 2nd ed. – Cambridge: Cambridge Univ. Press, 2010.
3. *Baier C., Katoen J.-P.* Principle of Model Checking. – The MIT Press, 2008. – 984 p.
4. *Grunskii I.S., Senchenko A.S.* Properties of systems of defining relations for automata// Discrete Mathematics and Applications. – 2004. – Vol. 14, Iss. 6. – P. 593–601.
5. *Grunskii I., Mikhaylova I., Sapunov S.* Domination on the vertices of labeled graphs// Algebra and Discrete Mathematics. – 2012. – Vol. 14, Iss. 2. – P. 174–184.
6. *Stepkin A.V.* Using a Collective of Agents for Exploration of Undirected Graphs// Cybernetics and System Analysis. – 2015. – Vol. 51, Iss. 2. – P. 223–233.
7. *Сапунов С.В., Сенченко О.С., Середа О.А.* Метричні властивості канонічної визначальної пари для детермінованих графів// Праці Інституту прикладної математики і механіки НАН України. – 2020. – Т. 34. – С. 134–145.

References

1. Okhotin, A. (2019) Graph-Walking Automata: From Whence They Come, and Whither They are Bound. In: *Hospodár M., Jirásková G. (eds) Implementation and Application of Automata. CIAA 2019. Lecture Notes in Computer Science, vol 11601.* Springer, Cham.
2. Dudek, G., Jenkin, M. (2010). *Computational Principles of Mobile Robotics,* 2nd ed. Cambridge, Cambridge Univ. Press.

3. Baier, C., Katoen, J.-P. (2008). *Principle of Model Checking. The MIT Press.*
4. Grunskii, I.S., Senchenko, A.S. (2004). Properties of systems of defining relations for automata. *Discrete Mathematics and Applications, 14*(6), 593–601.
5. Grunskii, I., Mikhaylova, I., Sapunov, S. (2012). Domination on the vertices of labeled graphs. *Algebra and Discrete Mathematics, 15*(2), 174–184.
6. Stepkin, A.V. (2015). Using a Collective of Agents for Exploration of Undirected Graphs. *Cybernetics and System Analysis, 51*(2), 223–233.
7. Sapunov, S.V., Senchenko, A.S., Sereda, O.A. (2020). Metrychni vlastyvosti kanonichnoyi vyznachalnoyi pary dlya determinovanykh grafiv. *Trudy IPMM NAN Ukrainy, 34*, 134–145.

A.S. Senchenko, M.I. Prytula, O.A. Sereda

Representation of deterministic graphs by defining pair of words.

The paper proposes a representation of deterministic graphs (D-graphs) by sets of words in the alphabet of labels of their vertices. Nowadays graphs are a conceptual tool for analytics, development and testing of various software. Among all graphs, there is a subclass of labeled graphs whose elements have labels from a predefined alphabet. Such graphs are actively used to describe and model computational processes in programming, robotics, model verification and validation, etc. Labeled graphs are an information environment for mobile agents, whose movement along the graph can be represented by sequences of vertex labels – words in the label alphabet. A vertex-labeled graph is said to be D-graph if all vertices in the neighborhood of every its vertex have different labels. For such graphs, in the case when its graph map (i.e., the set of vertices and edges and the labeling function) and the initial vertex from which the agents start their movements are known, there is an unambiguous correspondence between the sequence of labels of the vertices visited by the agent and the trajectory of this agent's movements in the graph. In the case when the map of the studied D-graph is unknown to an external observer, the movements of agents can be organized in such a way that, based on their analysis, the observer receives the desired information about the structure of the graph (for example, the map of the graph, the shortest paths between vertices, comparison of the studied graph with the reference graph). Such an analysis can significantly simplify the linguistic representation of a D-graph – the mapping of a graph to one or more finite sets of words in the alphabet of graph vertex labels, which can be used to reconstruct the graph. In this paper, we propose a representation of deterministic graphs by a defining pair of word sets, the first component of which describes the cycles of the graph, and the second - its leaf vertices. This representation is analogous to the system of defining relations for automata. We proposed the algorithm, that for any pair of sets, builds a D-graph for which this pair is deterministic, or reports that it is impossible to do so is given. The algorithm for constructing a canonical defining pair for a D-graph is also given and numerical estimates of this pair for D-graphs with a known number of vertices and edges are found. Further directions of research on this topic are outlined. The results will allow to use new methods and algorithms for solving problems of analysis of graphs with marked vertices.

Keywords: *deterministic graphs, representation, defining pair.*

Інститут прикладної математики і механіки НАН України, Слов'янськ
Senchenko.a76@gmail.com, elanir358@gmail.com, 0955174042@ukr.net

Отримано 19.12.22