УДК 519.713.1

А.М. ЧЕБОТАРЬОВ
Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: *ancheb@gmail.com*.

ПЕРЕТИН $-\omega$-РЕГУЛЯРНИХ ВИРАЗІВ

Анотація. Запропоновано метод побудови $-\omega$-регулярного виразу, що задає перетин множин $-\omega$-слів, поданих у вигляді $-\omega$-регулярних виразів R_1 і R_2. Така побудова здійснюється без переходу до ω-автоматів, тобто безпосереднім перетворенням виразу $R = R_1 \cap R_2$. Процес побудови $-\omega$-регулярного виразу, що задає перетин $R_1 \cap R_2$, подано у вигляді дерева перетинів, вершини якого відповідають перетинам простих $-\omega$-регулярних виразів, отриманих під час перетворення перетину $R_1 \cap R_2$. Побудоване дерево перетинів визначає систему лінійних рівнянь зі змінними, значеннями яких є множини $-\omega$-слів. Одна з цих змінних R відповідає множині $-\omega$-слів, що задається перетином $R_1 \cap R_2$, тобто вираз, який задає перетин $-\omega$-регулярних виразів R_1 і R_2, є значенням змінної R у розв'язку цієї системи лінійних рівнянь.

Ключові слова: зворотне надслово, $-\omega$-регулярний вираз, простий $-\omega$-регулярний вираз, $-\omega$-розгортка, *-розгортка, дерево перетинів.

ВСТУП

Регулярні ω-мови (інакше ω-регулярні мови або ω-регулярні множини) знаходять широке застосування у задачах специфікації, верифікації та синтезу реактивних систем. Ці мови, тобто множини нескінченних слів (ω-слів), замкнуті відносно операцій скінченного об'єднання, перетину та доповнення. Для їхнього представлення використовують ω-автомати, зокрема автомати Бюхі, Маллера та інші автомати-розпізнавачі, ω-регулярні вирази, а також для важливого підкласу цих мов, так званих беззірочних ω-регулярних мов, використовують логіки першого порядку і темпоральні логіки з лінійним часом, наприклад LTL. Серед перелічених засобів представлення найбільшого поширення набули ω-автомати і темпоральні логіки. Однією з причин меншого поширення ω-регулярних виразів є складність представлення ними перетину і доповнення ω-регулярних множин, заданих ω-регулярними виразами.

Позначимо $\mu(R)$ ω-регулярну множину, задану ω-регулярним виразом R. Для побудови ω-регулярного виразу, що задає множину $\mu(R_1) \cap \mu(R_2)$, можна діяти у такий спосіб. Оскільки методи переходу від ω-регулярного виразу R до автомата Бюхі $A(R)$, який розпізнає множину $\mu(R)$, достатньо відомі [1, 2], будується перетин автоматів $A(R_1) \cap A(R_2)$ [3, 4], тобто автомат Бюхі, що розпізнає множину $\mu(R_1) \cap \mu(R_2)$. Потім здійснюється перехід від цього автомата до ω-регулярного виразу, що задає ω-регулярну множину, яку розпізнає автомат. Ситуація з побудовою ω-регулярного виразу для доповнення ω-регулярної множини суттєво проблематичніша у зв'язку зі складністю побудови доповнення ω-автомата.

У цій роботі пропонується спосіб побудови ω-регулярного виразу, що визначає перетин множин, заданих ω-регулярними виразами R_1 і R_2, без використання ω-автоматів, тобто перетворенням виразу $R_1 \cap R_2$. Ця стаття мотивована попередніми працями [5, 6], що присвячені синтезу Σ-автомата. У [5] запропоновано метод синтезу Σ-автомата, специфікованого логічною мовою першого порядку LP. Синтез здійснюється шляхом перетворення формул мови LP вигляду $\forall t F(t)$ до вигляду $\forall t \bigvee_{i=1}^{n} F_i(t-1) f_i(t)$, де формули $F_i(t)$ відповідають станам автомата, що синтезується. Кожній формулі $F_i(t)$ ставиться у відповідність множина зворотних надслів ($-\omega$-слів) в алфавіті Σ, тобто нескінченних вліво слів вигляду $u = \ldots \sigma_{-2}\sigma_{-1}\sigma_0$, де $\sigma_i \in \Sigma$ $(i = 0, -1, -2, \ldots)$. Зворотне надслово u симетричне ω-слову $u^- = \sigma_0\sigma_1\sigma_2 \ldots$ Множина усіх зворотних надслів в алфавіті Σ позначається $\Sigma^{-\omega}$. Формула $F_i(t)$ задає підмножину множини $\Sigma^{-\omega}$, що називається $-\omega$-регулярною множиною, тому в цій роботі розглядаються саме такі множини. Ці множини подаються виразами, симетричними ω-регулярним виразам, наприклад, ω-регулярному виразу $b^* a \Sigma^\omega$ відповідає $-\omega$-регулярний вираз $\Sigma^{-\omega} ab^*$. Отже, усі перетворення $-\omega$-регулярних виразів симетрично застосовують для ω-регулярних виразів. Часто використання в алгоритмі синтезу замість формул $F_i(t)$ $-\omega$-регулярних виразів, які задають ті самі множини зворотних надслів, що і формули $F_i(t)$, дає змогу спростити процедуру синтезу.

У праці [6] розглядається задача знаходження фіктивних станів у синтезованому автоматі, яка зводиться до перевірки належності $-\omega$-слова множині $-\omega$-слів, що задана $-\omega$-регулярним виразом. Як у [5], так і в [6] використовуються операції перетину $-\omega$-регулярних множин. Ця робота присвячена побудові $-\omega$-регулярного виразу, що задає такий перетин.

ОСНОВНІ ПОНЯТТЯ

Нехай Σ — скінченний алфавіт, $\mathbf{Z}$ — множина цілих чисел, $\mathbf{N}^+ = \{z \in \mathbf{Z} \mid z \geq 0\}$, $\mathbf{N}^- = \{z \in \mathbf{Z} \mid z \leq 0\}$. Відображення r множини $\{1, \ldots, n\}$ $(n \geq 0)$ у Σ називається словом довжини n в алфавіті Σ і позначається $\sigma_1\sigma_2 \ldots \sigma_n$, де $\sigma_i = r(i)$ для усіх $1 \leq i \leq n$. Слово довжини 0 (порожнє слово) позначається ε. Відображення $l : \mathbf{N}^+ \to \Sigma$ називається надсловом (ω-словом) в алфавіті Σ і позначається $\sigma_0\sigma_1\sigma_2 \ldots$, де $\sigma_i = l(i)$, $i \in \mathbf{N}^+$. Відображення $g : \mathbf{N}^- \to \Sigma$ називається зворотним надсловом ($-\omega$-словом) і позначається $\ldots \sigma_{-2}\sigma_{-1}\sigma_0$, де $\sigma_i = g(i)$ для $i \in \mathbf{N}^-$. Множина усіх слів в алфавіті Σ, включаючи порожнє слово, позначається Σ^*, множина усіх надслів — Σ^ω, а множина усіх зворотних надслів — $\Sigma^{-\omega}$. На множині $\Sigma^* \cup \Sigma^{-\omega}$ у звичайний спосіб визначена часткова операція конкатенації $(\cdot)$, яку поширимо також на підмножини множин Σ^*, $\Sigma^{-\omega}$. Надалі символ конкатенації будемо випускати.

Визначимо над множинами слів операції скінченної ітерації $({}^*)$ і нескінченної ітерації $({}^{-\omega})$. Нехай $R \subseteq \Sigma^*$, операції ітерації визначаються у такий спосіб:

- $R^* = \{r_1 r_2 \ldots r_n \mid n \geq 0, r_i \in R\}$, зазначимо, що $n = 0$ відповідає порожньому слову ε;
- $R^{-\omega} = \{\ldots r_{-2} r_{-1} r_0 \mid$ для усіх $i \leq 0$ $r_i \in R \setminus \{\varepsilon\}\}$.

Для опису множин зворотних надслів використовуватимемо $-\omega$-регулярні вирази, які є скінченними об'єднаннями виразів вигляду $R^{-\omega}U$, де R і U — регулярні вирази, тобто вирази, побудовані із символів алфавіту Σ за допомогою операцій об'єднання, конкатенації та скінченної ітерації. Як видно, $-\omega$-регулярні вирази визначено симетрично ω-регулярним виразам, що використовують операцію нескінченної ітерації $({}^\omega)$. Множини зворотних надслів, що задаються $-\omega$-регулярними виразами, називаються $-\omega$-регулярними множинами. Використовуючи одні й ті самі символи для відповідних операцій над $-\omega$-регулярними множинами і $-\omega$-регулярними виразами, а також одні й ті самі символи для відповідних констант, а саме $\cup, \cdot, \cap, \varepsilon$ та $\varnothing$, отримуємо тотожну трансляцію мови $-\omega$-регулярних виразів у мову $-\omega$-регулярних множин. Тому далі $-\omega$-регулярні вирази асоціюються з $-\omega$-регулярними множинами, які вони задають. Зазначимо, що операції конкатенації відповідає добуток $-\omega$-регулярних множин. Щоб встановити відповідність між $-\omega$-регулярними множинами і $-\omega$-регулярними виразами, в описі одноелементної множини не будемо використовувати фігурні дужки, а скінченну множину вигляду $\{a_1, a_2, \ldots, a_n\}$ записуватимемо у вигляді $(a_1 \cup a_2 \cup \ldots \cup a_n)$. Під час тотожних перетворень таких виразів будемо використовувати усі властивості операцій над множинами, наприклад дистрибутивність перетину відносно об'єднання та інші. Зазначимо, що у $-\omega$-регулярних виразах відсутня операція, яка відповідає перетину множин, тому для отримання $-\omega$-регулярного виразу треба так тотожно перетворити вираз, що задає $-\omega$-регулярну множину, щоб він не містив операції перетину. Оскільки надалі розглядатимуться тільки $-\omega$-регулярні вирази, символ $-\omega$ випускатимемо, розуміючи під регулярним виразом $-\omega$-регулярний вираз. Також іноді будемо розглядати слова у розширеному алфавіті $\Sigma' = 2^{\Sigma}$, які задають множини слів, наприклад $a\Sigma$, $(a \cup b)cb$, $a\Sigma\Sigma b$. Тут Σ — об'єднання усіх символів алфавіту Σ. Вираз, який має вигляд $V^{-\omega}U$, де $V, U \subseteq \Sigma^*$, а U побудовано за допомогою операції конкатенації з виразів вигляду Q^* і $\sigma \in \Sigma'$, називатимемо простим регулярним виразом, а U — простим термом. Під час обчислення перетину двох простих регулярних виразів можуть виникати вирази вигляду $V^{-\omega}U$, де U являє собою об'єднання простих термів, яке назвемо складеним термом. З урахуванням дистрибутивності перетину відносно об'єднання перетин двох регулярних виразів зі складеними термами розкриттям дужок зводиться до декількох перетинів простих регулярних виразів, тому далі розглядатимемо перетин двох простих регулярних виразів. Такий перетин назвемо нормальною формою, якщо обидва терми у регулярних виразах, що перетинаються, мають вигляд $U\sigma$, де U — простий терм (можливо, ε), а $\sigma \in \Sigma'$.

Нехай $R_1, R_2 \subseteq \Sigma^{-\omega}$ — прості регулярні вирази в алфавіті Σ. Обчислення перетину $R_1 \cap R_2$ зводиться до обчислення перетинів, які є нормальними формами. Для цього виконуються такі операції:

- *-розгортка, тобто перетворення простого терма, який закінчується виразом вигляду Q^*, шляхом заміни Q^* виразом $(\varepsilon \cup Q^* Q)$;
- $-\omega$-розгортка, яка полягає в тім, що простий регулярний вираз $V^{-\omega}$, тобто терм якого дорівнює ε, замінюється виразом $V^{-\omega}V$;
- розкриття дужок, тобто представлення перетину регулярних виразів зі складеними термами у вигляді об'єднання перетинів простих регулярних виразів;
- операція елементарного перетину, яка застосовується до нормальних форм; результат елементарного перетину виразів вигляду $V_1^{-\omega}U_1\alpha$ і $V_2^{-\omega}U_2\beta$, де $\alpha, \beta \in \Sigma'$, дорівнює $\varnothing$, якщо $\alpha \cap \beta = \varnothing$, і $(V_1^{-\omega}U_1 \cap V_2^{-\omega}U_2)\gamma$, де $\gamma = \alpha \cap \beta$, у протилежному разі.

Нормальна форма, в якій $\alpha \cap \beta = \varnothing$, називається порожнім перетином.

ОБЧИСЛЕННЯ ПЕРЕТИНУ ПРОСТИХ $-$w-РЕГУЛЯРНИХ ВИРАЗІВ

Процес побудови регулярного виразу для перетину $R = R_1U \cap R_2W$, де $R_1 = V_1^{-\omega}$, $R_2 = V_2^{-\omega}$, зручно подавати у вигляді дерева перетинів, вершини якого мають вигляд $(R_1U' \cap R_2W')\gamma$, де $\gamma \in \Sigma^*$. У цьому виразі перетин $(R_1U' \cap R_2W')$ називається основною частиною вершини, а γ — закінченням.

Коренева вершина збігається з $R = R_1 U \cap R_2 W$, де $U, W \subseteq \Sigma^*$, а $\gamma = \varepsilon$. Вершина, що відповідає порожньому перетину, називається порожньою вершиною. Зазначимо, що у разі, якщо $R = V_1^{-\omega} \cap V_2^{-\omega}$, де $V_1, V_2 \neq \Sigma$, то цей вираз замінюється виразом $R = V_1^{-\omega} V_1 \cap V_2^{-\omega} V_2$. Перехід з однієї вершини в іншу здійснюється в результаті виконання описаних вище операцій над перетинами простих регулярних виразів.

Перехід із вершини q, основна частина якої є непорожньою нормальною формою, здійснюється у вершину, що є результатом виконання елементарного перетину простих регулярних виразів, які містяться в основній частині вершини q. При цьому довжина слова γ збільшується на одиницю. Наприклад, перехід із вершини $(R_1 c \cap R_2 bc)d$ здійснюється у вершину $(R_1 \cap R_2 b)cd$.

Переходи із вершини відсутні, якщо вона порожня, або є простим регулярним виразом, або її основна частина збігається з основною частиною вершини, яка їй передує у тому самому шляху дерева. Такі вершини назвемо фінальними, оскільки вони завершують процес побудови шляху в дереві. Отже, непорожні фінальні вершини можуть бути двох типів. Вершини, які є простими регулярними виразами, назвемо простими, а вершини іншого типу — циклічними, оскільки вони породжують нескінченне дерево, що містить циклічний шлях.

Переходи із вершини q вигляду $(V_1^{-\omega} U_1 \cap V_2^{-\omega} U_2)$, яка не є фінальною, здійснюються у вершини, основні частини яких отримуються як результат наведених нижче перетворень. Якщо $V_i \neq \Sigma$ $(i \in \{1, 2\})$ і відповідний терм $U_i = \varepsilon$, то виконується $-\omega$-розгортка виразу $V_i^{-\omega}$. Якщо є терми, що закінчуються виразами вигляду Q^*, то виконується *-розгортка цих виразів. Після цього розкривають дужки і вилучають порожні перетини. В усіх отриманих у такий спосіб вершинах значення γ залишається таким самим, як у вершини q. Дуги, що ведуть в усі ці вершини, позначаються одинарною стрілкою, а дуга, що веде у вершину, отриману внаслідок виконання операції елементарного перетину, — подвійною.

Отже, після виконання операції елементарного перетину слід перевіряти, чи не є отримана вершина циклічною фінальною вершиною. При цьому достатньо обмежитися порівнянням її з кореневою вершиною та вершинами у тому самому шляху, які отримані в результаті виконання елементарного перетину.

Вершина, що передує фінальній і має таку саму основну частину, називається базисною. Коренева вершина за означенням є базисною.

У різних шляхах дерева можуть бути фінальні вершини з однаковими основними частинами, що мають одну і ту саму базисну вершину. Для таких вершин вводять змінну Q, що набуває значення із множини $2^{\Sigma^{-\omega}}$, яка замінює основну частину цих фінальних вершин і відповідної базисної вершини. Для фінальних вершин з різними основними частинами вводять різні змінні. У результаті кожна циклічна фінальна вершина і відповідна базисна вершина набувають вигляду Qr, де Q — змінна, а r — слово в алфавіті Σ. Основна частина фінальної вершини, що збігається з кореневою вершиною, замінюється змінною R. Вочевидь, що кожний шлях у розглядуваному дереві має скінченну довжину і закінчується фінальною вершиною, бо кількість різних вершин у дереві скінченна. Якщо усі фінальні вершини у дереві перетинів є порожніми, то перетин простих регулярних виразів, якому відповідає це дерево, також порожній.

Після побудови дерева перетинів складається система лінійних рівнянь від змінних, введених для основних частин фінальних вершин. Для кожного піддерева, породженого базисною вершиною q зі змінною Q, записується рівняння вигляду $Q = C_1 \cup \ldots \cup C_m \cup Q_1 r_1 \cup \ldots \cup Q_n r_n$, де C_i $(i \in \{1, \ldots, m\})$ — досяжна із q проста фінальна вершина, шлях до якої із вершини q не має ніяких інших базисних вершин, Q_i $(i \in \{1, \ldots, n\})$ — змінна, що відповідає досяжній із q базисній або фінальній вершині q_i, шлях до якої із q не містить ніякої іншої базисної вершини, а r_i — префікс закінчення вершини q_i, який відрізняє її від базисної вершини q. Отже, якщо закінчення вершини q дорівнює r, а вершини q_i — $r_1 r$, то $r_i = r_1$. Якщо $Q = R$, то r_i збігається з закінченням вершини q_i.

Розв'язання цієї системи рівнянь здійснюється методом підстановки і, зрештою, зводиться до розв'язання рівнянь вигляду $R = R_1 \cup RW$ або $R = RW$ з одною змінною R, де $R, R_1 \subseteq \Sigma^{-\omega}$, $W \subseteq \Sigma^*$. Як показано в [7], перше рівняння має максимальний розв'язок $R = R_1 W^* \cup W^{-\omega}$, якщо $W^{-\omega} \subseteq R$, і $R = R_1 W^*$ у протилежному разі, що легко перевіряється підстановкою розв'язку в рівняння; максимальний розв'язок другого рівняння має вигляд $R = W^{-\omega}$, якщо $W^{-\omega} \subseteq R$, і дорівнює $\varnothing$ у протилежному випадку. Якщо $R = RW_1 \cup RW_2$ і $W_1^{-\omega} \subseteq R$, а $W_2^{-\omega} \not\subset R$, то максимальний розв'язок має вигляд $R = (W_1 (W_2)^*)^{-\omega}$. Якщо $R = R_1 \cup RW_1 \cup RW_2$ і $W_1^{-\omega} \subseteq R$, а $W_2^{-\omega} \not\subset R$, то в цьому разі $R = R_1 (W_1 \cup W_2)^* \cup (W_1 (W_2)^*)^{-\omega}$. Отже, під час розв'язання системи рівнянь потрібно перевіряти істинність включення вигляду $W^{-\omega} \subseteq R$.

Задачі, що виникають під час обчислення перетину простих регулярних виразів, дещо розрізняються залежно від вигляду цих виразів. Будемо розрізняти два види простих регулярних виразів: $\Sigma^{-\omega} U$ і $V^{-\omega} U$, де $V, U \subseteq \Sigma^*$. Відповідно розглянемо такі випадки перетинів: 1) $R = V_1^{-\omega} U_1 \cap V_2^{-\omega} U_2$; 2) $R = \Sigma^{-\omega} U_1 \cap \Sigma^{-\omega} U_2$. Другий випадок відрізняється від першого тим, що у процесі побудови дерева перетинів операція $-\omega$-розгортки не використовується і операція елементарного перетину має деяку специфіку; наприклад, для $R \subseteq \Sigma^{-\omega}$ і $\sigma \in \Sigma^*$ перетин $R\sigma \cap \Sigma^{-\omega}\sigma$ дорівнює $R\sigma$, що відповідає простій фінальній вершині. Деякі інші особливості обчислень у цьому разі будуть розглянуті нижче.

Перетин вигляду $\boldsymbol{R = V_1^{-\mathrm{w}} U_1 \cap V_2^{-\mathrm{w}} U_2}$. Одна із задач, пов'язана з розв'язанням системи рівнянь, полягає у перевірці істинності включення вигляду $W^{-\omega} \subseteq Q$. Із вигляду вихідних рівнянь випливає, що $W \in \Sigma^*$. Наведене нижче твердження описує ситуацію, коли це включення виконується для вихідних рівнянь, тобто перевіряти його не потрібно.

Нехай базисна вершина має вигляд $(V_1^{-\omega} U_1 \cap V_2^{-\omega} U_2) r$, а фінальна — $(V_1^{-\omega} U_1 \cap V_2^{-\omega} U_2) \gamma r$.

Твердження 1. Якщо U_i $(i \in \{1, 2\})$ не містить операції *, то включення $\gamma^{-\omega} \subseteq V_i^{-\omega} U_i$ виконується.

Доведення. Нехай U_i $(i \in \{1, 2\})$ не містить операції *, тобто $U_i \in (\Sigma')^*$. Розглянемо, як змінюється цей терм на шляху із базисної вершини у фінальну. Фінальна вершина утворюється в результаті виконання операції елементарного перетину. Отже, терм U_i на шляху із базисної вершини у фінальну має змінитися так, щоб вилучення з нього останнього символу перетворювало його в U_i. Звідси випливає, що операції елементарного перетину має передувати $-\omega$-розгортка виразу $V_i^{-\omega}$, яка виконується, якщо терм при $V_i^{-\omega}$ дорівнює ε. Зазначимо, що додання символів у терм може здійснюватися також операцією *-розгортки, однак такий терм містить вираз вигляду Q^*. Оскільки вихідний терм U_i не містить *, такий вираз може з'явитися лише в результаті виконання $-\omega$-розгортки. Отже, на шляху із базисної вершини у фінальну один або декілька разів трапиться вершина з термом при $V_i^{-\omega}$, який дорівнює ε. Закінчення такої вершини, найближчої до фінальної, має вигляд $\alpha U_i r$, де $\alpha \in (V_i)^*$, і утворюється в результаті однієї або більше $-\omega$-розгорток $V_i^{-\omega}$, операцій елементарного перетину і, можливо, *-розгорток, які вилучають із терма вирази вигляду Q^*. Якщо у фінальній вершині $U_i \neq \varepsilon$, то потрібно виконати останню $-\omega$-розгортку $V_i^{-\omega}$, у результаті чого $V_i^{-\omega}$ буде замінено на $V_i^{-\omega} U_i r_1$, де $U_i r_1 \subseteq V_i$. Якщо r_1 не містить *, то усі символи із r_1 переносяться у закінчення. Як наслідок, закінчення фінальної вершини має вигляд $r_1 \alpha U_i r$. Покажемо, що $(r_1 \alpha U_i)^{-\omega} \subseteq V_i^{-\omega} U_i$. Зазначимо, що $(r_1 \alpha U_i)^{-\omega} = (U_i r_1 \alpha)^{-\omega} U_i$ і $U_i r_1 \alpha \subseteq (V_i)^*$. Отже, маємо $(U_i r_1 \alpha)^{-\omega} \subseteq$ $\subseteq ((V_i)^*)^{-\omega} = V_i^{-\omega}$ і відповідно $(r_1 \alpha U_i)^{-\omega} \subseteq V_i^{-\omega} U_i$.

Якщо під час виконання останньої $-\omega$-розгортки $V_i^{-\omega}$ замінюється виразом $V_i^{-\omega} U_i r'$, де r' містить вираз вигляду Q^*, то на відрізку шляху до фінальної вершини в результаті виконання *-розгорток цих виразів r' буде замінено словом r_1 таким, що $U_i r_1 \subseteq V_i$. Якщо $U_i = \varepsilon$, то після виконання одного або більше разів $-\omega$-розгорток $V_i^{-\omega}$ закінчення фінальної вершини матиме вигляд $r_1 \alpha r$, де $V_i^{-\omega} r_1$ замінює $V_i^{-\omega}$ в результаті останньої розгортки і $r_1 \alpha \subseteq (V_i)^*$, а значить, $(r_1 \alpha)^{-\omega} \subseteq ((V_i)^*)^{-\omega} = V_i^{-\omega}$. Кінець доведення.

Отже, відсутність операції * у термах U_1 і U_2 фінальної вершини — достатня умова істинності відповідного включення. Нескладно показати, що якщо $(\gamma_1)^{-\omega}, (\gamma_2)^{-\omega} \in Q$, то $(\gamma_1 \cup \gamma_2)^{-\omega} \in Q$, це відповідає паралельним циклам у вершині. Однак у загальному випадку таке твердження хибне.

Якщо U_i містить операцію *, то часто перевірка включення $\gamma^{-\omega} \subseteq V_i^{-\omega} U_i$ також виконується тривіально. Наприклад, якщо U_i містить символ, відсутній в γ, або закінчується символом, відмінним від того, яким закінчується γ, то включення $\gamma^{-\omega} \subseteq V_i^{-\omega} U_i$ не виконується.

Під час розв'язання рівнянь можуть виникати ситуації, коли коефіцієнт W при змінній Q в рівнянні для Q містить вираз вигляду S^*. Нескладно показати, що включення $W^{-\omega} \subseteq Q$ виконується тоді й тільки тоді, коли виконується включення $w^{-\omega} \subseteq Q$, де w утворюється внаслідок вилучення із W усіх виразів вигляду S^*. Отже, перевірка включень під час розв'язання рівнянь зводиться до перевірки включень вигляду $r^{-\omega} \subseteq Q$, де r — коефіцієнт при Q у відповідному вихідному рівнянні.

Приклад 1. Нехай $R = R_1 \cap R_2$, де $R_1 = (ab \cup c)^{-\omega}$, $R_2 = (a \cup bc^*)^{-\omega}$.

Дерево перетинів для $R_1 \cap R_2$ наведено на рис. 1.

Перетин $R_1 \cap R_2 = R_1 (ab \cup c) \cap R_2 (a \cup bc^*)$, що відповідає $-\omega$-розгорткам R_1 і R_2. Виконавши *-розгортку c^*, матимемо $R_1 (ab \cup c) \cap R_2 (a \cup b \cup bc^* c)$. Після розкриття дужок отримаємо шість перетинів, з яких тільки $(R_1 ab \cap R_2 b)$ і $(R_1 c \cap R_2 bc^* c)$ не є порожніми. Застосувавши до отриманих вершин операцію елементарного перетину, одержимо вершини $(R_1 a \cap R_2) b$ і $(R_1 \cap R_2 bc^*) c$. У дереві перетинів, наведеному на рис. 1, вершини, в які є переходи із вершини $(R_1 \cap R_2 bc^*) c$, утворюються в результаті *-розгортки c^*, розкриття дужок та вилучення порожніх вершин, що на дереві перетинів не відображені. У дереві є дві базисні вершини: $R_1 \cap R_2$ і $(R_1 \cap R_2 bc^*) c$. Перша з них коренева, і відповідні фінальні вершини мають вигляд Rab і $Rabc$. Для другої введена змінна Q,

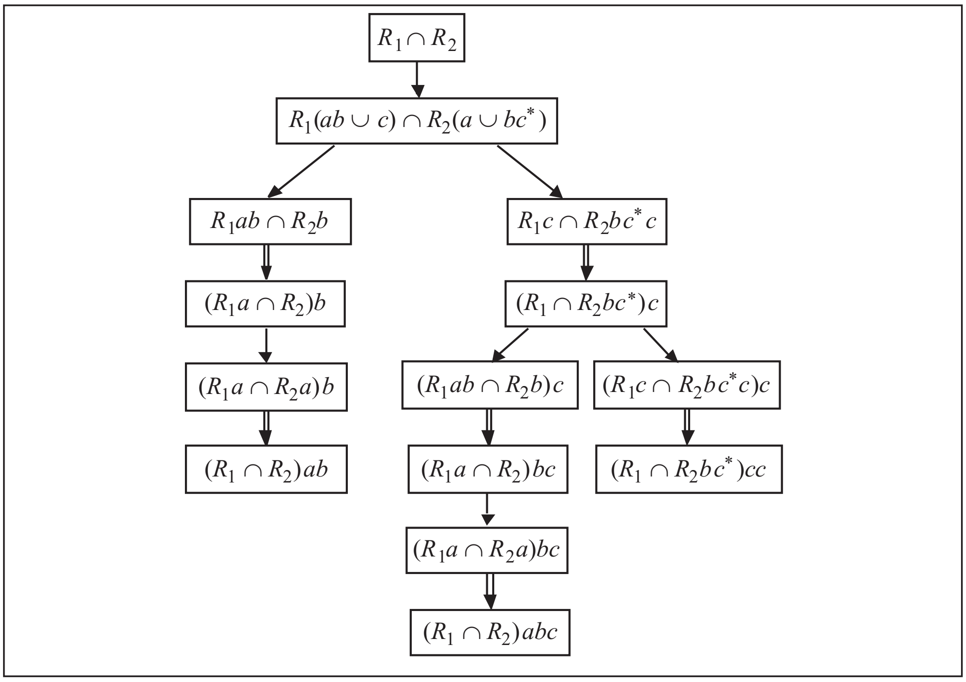

Рис. 1

внаслідок чого вона набуває вигляду Qc, а відповідна фінальна вершина — Qcc. Отже, дерево перетинів визначає два рівняння: $R = Rab \cup Qc$ і $Q = Rab \cup Qc$. Очевидно, що $c^{-\omega} \notin Q$, тому розв'язок другого рівняння має вигляд $Q = Rabc^*$. Підставивши цей розв'язок у рівняння для R, отримаємо $R = Rab \cup Rabc^* = Rabc^*$. Згідно з твердженням 1 $(abc)^{-\omega} \subseteq R$, отже, $(abc^*) \subseteq R$, тому $R = (abc^*)^{-\omega}$.

Перетин вигляду $\boldsymbol{R = \Sigma^{-\omega} U_1 \cap \Sigma^{-\omega} U_2}$. Перетин вигляду $\Sigma^{-\omega}U_1 \cap \Sigma^{-\omega}U_2$ є окремим випадком перетину вигляду $V_1^{-\omega}U_1 \cap V_2^{-\omega}U_2$, якщо $V_1 = V_2 = \Sigma$. Побудова дерева перетинів у такому разі здійснюється простіше, ніж для перетину загального вигляду, адже не треба виконувати операції $-\omega$-розгорток. Крім того, під час побудови дерева перетинів тут є більше можливостей для використання тотожних співвідношень для простих регулярних виразів вигляду $\Sigma^{-\omega}U$. Часто бувають корисними такі тотожні співвідношення:

1) $\Sigma^{-\omega}U \cap \Sigma^{-\omega} = \Sigma^{-\omega}U$;

2) $\Sigma^{-\omega}R_1^* R_2 = \Sigma^{-\omega}R_2$;

3) $\Sigma^{-\omega}R \cup (\Sigma^{-\omega}R_1 \cap \Sigma^{-\omega}R) = \Sigma^{-\omega}R$;

4) $\Sigma^{-\omega}(R_1 \cup Q^*) \cap \Sigma^{-\omega}R = \Sigma^{-\omega}R$;

5) $\Sigma^{-\omega}UR \cap \Sigma^{-\omega}R = \Sigma^{-\omega}UR$.

Отже, з вершини вигляду $\Sigma^{-\omega}U \cap \Sigma^{-\omega}$ здійснюється перехід у просту фінальну вершину $\Sigma^{-\omega}U$, а з вершини вигляду $\Sigma^{-\omega}UR \cap \Sigma^{-\omega}R$ — у фінальну вершину $\Sigma^{-\omega}UR$. Якщо $R \in \Sigma^*$, то цей результат можна отримати виконанням послідовності елементарних перетинів і застосуванням співвідношення 1. Якщо $R \subseteq \Sigma^*$, то, щоб його отримати без застосування співвідношення 5, можливо, потрібно побудувати із цієї вершини дерево перетинів з декількома фінальними вершинами, а також розв'язати рівняння, що визначаються цим деревом.

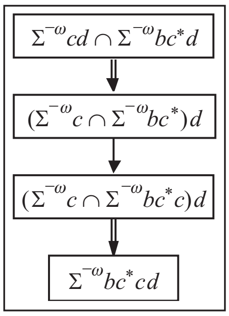

Рис. 2

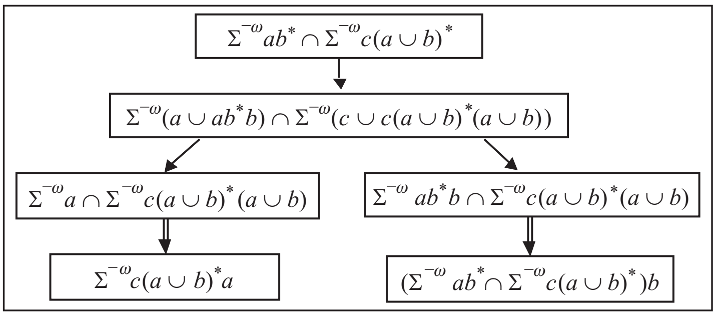

Рис. 3

Приклад 2. Нехай $R = \Sigma^{-\omega}(abc^* \cup bd)^* cd \cap \Sigma^{-\omega} bc^* d$.

Використавши співвідношення 2, одержимо $R = \Sigma^{-\omega} cd \cap \Sigma^{-\omega} bc^* d$. Відповідне дерево перетинів наведено на рис. 2.

Отже, $R = \Sigma^{-\omega} bc^* cd$.

Приклад 3. Нехай $R = \Sigma^{-\omega} ab^* \cap \Sigma^{-\omega} c(a \cup b)^*$.

Побудуємо для R дерево перетинів (рис. 3).

Перехід із кореневої вершини відповідає виконанню *-розгорток. Після цього виконується розкриття дужок, що визначає дві непорожні вершини. Фінальні вершини утворюються в результаті виконання елементарних перетинів. Ліва фінальна вершина проста, а права циклічна, для якої коренева вершина є базисною. Отже, одержуємо рівняння $R = \Sigma^{-\omega} c(a \cup b)^* a \cup Rb$. Оскільки $b^{-\omega} \notin R$, маємо розв'язок $R = \Sigma^{-\omega} c(a \cup b)^* ab^*$.

СПРОЩЕННЯ ДЕРЕВА ПЕРЕТИНІВ

Можливі два види спрощень дерева перетинів, тобто зменшення кількості вершин у дереві. Перший пов'язаний зі спрощенням вихідного виразу для розглядуваного перетину. Наприклад, якщо вираз $V^{-\omega}$ має вигляд $(R^*)^{-\omega}$, то його слід замінити виразом $R^{-\omega}$, що унеможливлює виконання *-розгорток. Крім того, до простого регулярного виразу вигляду $\Sigma^{-\omega} U$ можна застосовувати тотожні співвідношення, наприклад 2 (див. приклад 2) або 4. Останнє полягає у вилученні частини перетину, що має вигляд $\Sigma^{-\omega}(R \cup Q^*)$. Якщо в одній з частин перетину вигляду $V_1^{-\omega} U_1 \cap V_2^{-\omega} U_2$ є символ, відсутній у другій, то цю частину треба перетворити таким чином, щоб з множини зворотних надслів, яку вона задає, були вилучені усі ті й лише ті зворотні надслова, які містять цей символ. Це достатньо просте перетворення. Наприклад, розглянемо перетин $((abc^* \cup bd)^* cd)^{-\omega} \cap (bc^* d)^{-\omega}$. Ліва частина перетину не містить символа a, тому її слід замінити виразом $((bd)^* cd)^{-\omega}$. Якщо ліва частина перетину, з такою самою правою частиною, має вигляд $((ba^* c \cup bd)^* cd)^{-\omega}$, то вона замінюється виразом $((bc \cup bd)^* cd)^{-\omega}$. Для перетину вигляду $\Sigma^{-\omega} U_1 \cap \Sigma^{-\omega} U_2$ аналогічний спосіб спрощення не можна застосувати, оскільки з того, що деякий символ відсутній у виразі $\Sigma^{-\omega} U$, не випливає, що цей вираз задає множину зворотних надслів, яка не містить цього символа.

Якщо вираз $V^{-\omega}$ має вигляд $(R \cup Q^*)^{-\omega}$, то під час побудови дерева перетинів можуть з'явитися однакові вершини в одному й тому шляху дерева. При цьому фінальна вершина збігається з базисною, що не додає ніякої інформації у відповідне рівняння. Наприклад, якщо у перетині $R_1 \cap R_2$ маємо $R_1 = (ab \cup c^*)^{-\omega}$,

$R_2 = (a \cup bc^*)^{-\omega}$, то після $-\omega$-розгорток і розкриття дужок разом з іншими двома вершинами буде одержано вершину $R_1 c^* \cap R_2 a$. Після *-розгортки c^* отримаємо вершину $R_1 \cap R_2 a$, застосувавши до якої $-\omega$-розгортку R_1, знов отримаємо вершину $R_1 c^* \cap R_2 a$. Для унеможливлення такої ситуації вираз Q^* треба замінити виразом $Q^* Q$ відповідно до тотожності $(R \cup Q^*)^{-\omega} = (R \cup Q^* Q)^{-\omega}$. Легко переконатися, що у розглядуваному прикладі, якщо c^* у R_1 замінити виразом $c^* c$, то повторення вершини не виникає. Другий вид спрощення пов'язаний з вилученням деяких вершин у процесі побудови дерева перетинів.

Нехай маємо перетини $(A \cap B)$ і $(C \cap D)$, де $A, B, C, D \subseteq \Sigma^{-\omega}$ і $A \subseteq C$, $B \subseteq D$, тоді $(A \cap B) \subseteq (C \cap D)$. У цьому разі будемо казати, що $(C \cap D)$ поглинає $(A \cap B)$. Наприклад, вершина $(R_1 \cap R_2 bc^*)$ поглинає вершину $(R_1 \cap R_2 b)$. Зазначимо, що тут істинність включення $R_2 b \subseteq R_2 bc^*$ не залежить від вигляду R_2. Також і вершина $(V_1^{-\omega} \cap V_2^{-\omega} bc^*)$ поглинає вершину $(V_1^{-\omega} c^* c \cap V_2^{-\omega} bc^*)$, якщо $c^* c \subseteq V_1^*$. Якщо переходи у вершини q_1 і q_2 здійснюються з одної і тієї самої вершини і $q_1 \subseteq q_2$, то вершину q_1 можна вилучити. Така ситуація часто виникає під час виконання *-розгортки. Наприклад, нехай у перетині $R_1 \cap R_2$ маємо $R_1 = ((a \cup bc^*)^* ab^*)^{-\omega}$ і $R_2 = (ab \cup c)^{-\omega}$. У процесі побудови дерева перетинів буде отримано вершину $(R_1 (a \cup bc^*)^* \cap R_2 a)\, babb$. Після виконання *-розгортки терма $(a \cup bc^*)^*$, розкриття дужок і вилучення порожньої вершини отримаємо вершини $(R_1 \cap R_2 a) babb$ і $(R_1 (a \cup bc^*)^* a \cap R_2 a) babb$. Неважко помітити, що $R_1 (a \cup bc^*)^* a \subseteq R_1$, тому вершину $(R_1 (a \cup bc^*)^* a \cap R_2 a) babb$ можна вилучити, а разом з нею і все піддерево, яке вона породжує.

ВИСНОВКИ

Для подання функціональних властивостей реактивних систем у задачах специфікації та верифікації поряд з темпоральними логіками зручно використовувати ω-регулярні вирази, виразні можливості яких перевищують такі можливості темпоральних логік першого порядку. Щоб зробити ω-регулярні вирази більш компактними і легшими для сприйняття користувачем, їх розширюють додатковими операціями, такими як перетин та доповнення. Хоча це зручно для специфікації властивостей, але у задачах синтезу на основі ω-регулярних ($-\omega$-регулярних) виразів таке розширення спричиняє суттєві ускладнення процедури синтезу, що пов'язане з необхідністю трансляції розширених ω-регулярних виразів у звичайні.

У роботі розглядається трансляція $-\omega$-регулярного виразу, в якому використовується операція перетину, у звичайний $-\omega$-регулярний вираз. Таку трансляцію здійснюють безпосереднім перетворенням перетину $-\omega$-регулярних виразів у еквівалентний йому $-\omega$-регулярний вираз, тобто без перетворення $-\omega$-регулярних виразів в автомати над нескінченними словами. Процес такого перетворення подається у вигляді скінченного дерева перетинів, яке визначає систему лінійних рівнянь, від змінних, що набувають значення із множини $-\omega$-регулярних виразів, над алфавітом Σ. Розв'язання цієї системи рівнянь здійснюється методом підстановки, що, у свою чергу, зводиться до розв'язання лінійних рівнянь вигляду $Q = R_1 \cup QW$ або $Q = QW$ з одним невідомим Q. Результат розв'язання такого рівняння для невідомого Q залежить від виконуваності включень вигляду $r^{-\omega} \subseteq \mu(Q)$, де r — слово із Σ^*, яке є коефіцієнтом при невідомому, а $\mu(Q)$ — множина зворотних надслів, що відповідає невідомому Q. Доведено твердження, яке дає змогу спростити перевірку виконуваності подібних включень. Слід зазначити, що рівняння, одержані в результаті побудови дерева перетинів, визначають граф, що задає $-\omega$-регулярний вираз для розглядуваного перетину $-\omega$-регулярних виразів. Вершини цього графу відповідають невідомим, а дуги, що ведуть у вершину Q, визначаються рівнянням для Q і відзначаються коефіцієнтами при невідомих, що відповідають тим вершинам, з яких вони виходять. Однак від такого графу необхідно перейти до $-\omega$-регулярного виразу, який він задає, і при цьому треба враховувати виконуваність включень, про які йшлося вище. Тому цей підхід, що не потребує розв'язання системи рівнянь, тут не розглядався.

У статті сформульовано деякі властивості графу перетинів, що дають змогу спростити його побудову, тобто процес перетворення перетину простих $-\omega$-регулярних виразів у еквівалентний йому $-\omega$-регулярний вираз. Визначення складності такого процесу — це окрема проблема, яку достатньо досліджено для регулярних виразів, але для $-\omega$-регулярних виразів вона значно складніша. Як вже згадувалося, усі результати, одержані для $-\omega$-регулярних виразів, симетрично переносяться на ω-регулярні вирази.

СПИСОК ЛІТЕРАТУРИ

1. Thiemann P., Sulzmann M. From ω-regular expressions to Buchi automata via partial derivatives. *Lecture Notes Comput. Sci.* Berlin; Heidelberg: Springer-Verlag, 2015. Vol. 8977. P. 287–298.
2. Loding C., Tollkotter A. Transformations between regular expressions and ω-automata. *Proc. 41th Int. Symp. on Mathematical Foundation of Computer Science* (MFCS 2016, Aug 22–26, Krakow, Poland). Aachen, Germany: Dagstuhl Publishing, 2016. P. 88:1–88:13.
3. Mukund M. Finite state automata on infinite inputs. *Modern Application of Automata Theory*. D'Souza D., Shankar P. (eds.). Singapore: World Scientific Press, 2012. P. 45–78.
4. Thomas W. Automata on infinite objects. *Handbook of Theoretical Comput. Science*. Leeuwen J. van (ed.). Amsterdam: Elsevier Science Publishers, 1990. P. 134–191.
5. Чеботарев А.Н. Синтез Σ-автоматов, специфицированных в логических языках LP и LF первого порядка. *Кибернетика и системный анализ.* 2018. Т. 54, № 4. С. 16–31.
6. Чеботарев А.Н. Определение фиктивных состояний в автомате, синтезированном по спецификации в языке LP. *Кибернетика и системный анализ.* 2019. Т. 55, № 5. С. 47–57.
7. Staiger L. On ω-power languages. *New Trends in Formal Languages. LNCS*. Berlin; Heidelberg: Springer-Verlag, 1997. Vol. 1218. P. 377–394.

A.N. Chebotarev

INTERSECTION OF $-\omega$-REGULAR EXPRESSIONS

Abstract. A method is proposed for constructing the $-\omega$-regular expression that defines the intersection of sets of $-\omega$-words represented by $-\omega$-regular expressions R_1 and R_2. Such constructing is carried out without passing to ω-automata, that is by direct transformation of the expression $R = R_1 \cap R_2$. The process of constructing $-\omega$-regular expression defining the intersection $R_1 \cap R_2$ is represented in the form of an intersection tree whose vertices correspond to intersections of simple $-\omega$-regular expressions obtained during the transformation of the intersection $R_1 \cap R_2$. The tree constructed in this way defines a system of linear equations with variables whose values are sets of $-\omega$-words. One of these variables R corresponds to the set of $-\omega$-words defined by $R_1 \cap R_2$. The $-\omega$-regular expression defining the intersection $R_1 \cap R_2$ is the value of the variable R in the solution for this linear equation system.

Keywords: left-infinite word ($-\omega$-word), $-\omega$-regular expression, simple $-\omega$-regular expression, $-\omega$-expansion, *-expansion, intersection tree.

Надійшла до редакції 24.02.2021