Зміст

5 Регулярні множини і регулярні вирази

5.1 Регулярні множини

Нехай Σ — скінчений алфавіт. *Регулярна множина* в алфавіті Σ визначається рекурсивно:

1. $\varnothing$ — пуста множина — це регулярна множина в алфавіті Σ;
2. $\{\varepsilon\}$ — пусте слово — регулярна множина в алфавіті Σ;
3. $\{a\}$ — однолітерна множина — регулярна множина в алфавіті Σ;
4. Якщо P та Q — регулярні множини, то такими є наступні множини:
   - $P \cup Q$ (операція об'єднання);
   - $P \times Q$ (операція конкатенації);
   - $P^\star = \{\varepsilon\} \cup P \cup P^2 \cup \dots$ (операція ітерації).
5. Ніякі інші множини, окрім побудованих на основі 1–4 не є регулярними множинами.

Таким чином, регулярні множини можна побудувати з базових елементів 1–3 шляхом скінченого застосування операцій об'єднання, конкатенації та ітерації.

5.2 Регулярні вирази

Регулярні вирази позначають регулярні множини таким чином, що:

1. 0 позначає регулярну множину $\varnothing$;
2. ε позначає регулярну множину $\{\varepsilon\}$;

3. a позначає регулярну множину $\{a\}$;
4. Якщо p та q позначають відповідно регулярні множини P та Q, то
   - $p + q$ позначає регулярну множину $P \cup Q$;
   - $p \cdot q$ позначає регулярну множину $P \times Q$;
   - $p^\star$ позначає регулярну множину $P^\star$.
5. Ніякі інші вирази, окрім побудованих на основі 1–4 не є регулярними виразами.

5.2.1 Алгебра регулярних виразів

Оскільки ми почали вести мову про вирази, нам зручніше перейти до поняття алгебри регулярних виразів. Для кожної алгебри одним з важливих питань є питання еквівалентних перетворень, які виконуються на основі тотожностей у цій алгебрі. Сформулюємо основні тотожності алгебри регулярних виразів:

1. $a + b + c = a + (b + c)$ (ліва асоціативність додавання);
2. $a + b + c = (a + b) + c$ (права асоціативність додавання);
3. $a + 0 = 0 + a = a$ (0 — нейтральний елемент за додаванням);
4. $a \cdot b \cdot c = a \cdot (b \cdot c)$ (ліва асоціативність множення);
5. $a \cdot b \cdot c = (a \cdot b) \cdot c$ (права асоціативність множення);
6. $a + b = b + a$ (комутативність додавання);
7. $a \cdot \varepsilon = \varepsilon \cdot a = a$ (ε — нейтральний елемент за множенням);
8. $a \cdot 0 = 0 \cdot a = 0$ (0 — нульовий елемент за множенням);
9. $a \cdot (b + c) = a \cdot b + a \cdot c$ (ліва дистрибутивність множення відносно додавання);
10. $(a + b) \cdot c = a \cdot c + b \cdot c$ (права дистрибутивність множення відносно додавання).

У алгебри регулярних виразів є і некласичні властивості:

1. $a + a = a$;
2. $p + p^\star = p^\star$;

3. $0^\star = \varepsilon$;

4. $\varepsilon^\star = \varepsilon$.

5.2.2 Лінійні рівняння

За аналогією з класичними алгебрами розглянемо лінійне рівняння в алгебрі регулярних виразів: $X = a \cdot X + b$, де a, b — регулярні вирази.

Взагалі кажучи таке рівняння (в залежності від a та b) може мати безліч розв'язків.

Серед всіх розв'язків рівняння з регулярними коефіцієнтами виберемо найменший розв'язок $X = a^\star \cdot b$, який назвемо *найменша нерухома точка.*

Щоб перевірити, що $a^\star \cdot b$ справді розв'язок рівняння в алгебрі регулярних виразів, підставимо його в початкове рівняння та перевіримо тотожність виразів на основі системи тотожних перетворень:

$$a^\star b = aa^\star b = (aa^\star + \varepsilon)b = (a(\varepsilon + a + a^2 + \ldots) + \varepsilon)b = (\varepsilon a + a^2 + \ldots)b = a^\star b.$$

5.2.3 Системи рівнянь

В алгебрі регулярних виразів також розглядають і системи лінійних рівнянь з регулярними коефіцієнтами:

$$\begin{cases} X_1 = a_{11}X_1 + a_{12}X_2 + \ldots + a_{1n}X_n + b_1, \\ X_2 = a_{21}X_1 + a_{22}X_2 + \ldots + a_{2n}X_n + b_2, \\ \qquad \ldots \\ X_n = a_{n1}X_1 + a_{n2}X_2 + \ldots + a_{nn}X_n + b_n. \end{cases}$$

Метод розв'язування системи лінійних рівнянь з регулярними коефіцієнтами нагадує метод виключення Гауса.

1. Для $i = \overline{1..n}$, використавши систему тотожних перетворень, записати i-е рівняння у вигляді: $X_i = aX_i + b$, де a — регулярний вираз в алфавіті Σ, а b — регулярний вираз виду

   $$\beta_0 + \beta_{i+1}X_{i+1} + \beta_{i+2}X_{i+2} + \ldots + \beta_n X_n,$$

   де β_k $(k = 0, \overline{i+1..n})$ — регулярні коефіцієнти. Далі, в правих частинах рівнянь зі змінними $X_{i+1}, X_{i+2}, \ldots, X_n$ в лівій частині рівняння підставити замість X_i значення $a^\star b$.

2. Для $i = \overline{n..1}$ розв'язати i-е рівняння яке зараз має вигляд $X_i = aX_i + b$, де a, b — регулярні вирази в алфавіті Σ, і підставити його розв'язок $a^\star b$ у коефіцієнти рівнянь зі змінними $X_{i-1}, X_{i-2}, \ldots, X_1$.

Приклад. Розв'язати систему 2×2:

$$\begin{cases} X_1 = a_{11}X_1 + a_{12}X_2 + b_1, \\ X_2 = a_{21}X_1 + a_{22}X_2 + b_2. \end{cases}$$

Розв'язок:

1. З першого рівняння

$$X_1 = a_{11}^\star(a_{12}X_2 + b_1).$$

2. Підставляємо у друге рівняння, воно набуває вигляду

$$X_2 = a_{21}(a_{11}^\star(a_{12}X_2 + b_1)) + a_{22}X_2 + b_2.$$

   Або, після спрощень:

$$X_2 = (a_{21}a_{11}^\star a_{12} + a_{22})X_2 + (a_{21}a_{11}^\star b_1 + b_2).$$

   Звідки знаходимо:

$$X_2 = (a_{21}a_{11}^\star a_{12} + a_{22})^\star(a_{21}a_{11}^\star b_1 + b_2).$$

3. Підставляємо у вираз для X_1:

$$X_1 = a_{11}^\star(a_{12}(a_{21}a_{11}^\star a_{12} + a_{22})^\star(a_{21}a_{11}^\star b_1 + b_2) + b_1).$$

   Тут можна розкрити дужки, але зараз у цьому немає потреби.

5.3 Контрольні запитання

1. Яка множина називається регулярною (в алфавіті Σ)?
2. Як регулярні вирази позначаються регулярні множини?
3. Які класичні тотожності алгебри регулярних виразів ви знаєте?
4. Які не класичні тотожності алгебри регулярних виразів ви знаєте?

5. Який розв'язок лінійного рівняння $X = a \cdot X + b$ називається найменшою нерухомою точкою?

6. Доведіть, що згаданий у попередньому рівняння вираз справді є розв'язком.

7. Сформулюйте алгоритм Гауса розв'язування систем лінійних регулярних рівнянь.

8. Розв'яжіть систему 3×3:

$$\begin{cases} X_1 = aX_1 + bX_2 + c, \\ X_2 = cX_1 + aX_2 + bX_3, \\ X_3 = b + cX_2 + aX_3. \end{cases}$$