BI-AAG cvičení 5 - Seznam pojmů

Bc. Eliška Šestáková

23.10.2014

Regulární výraz – Regulární výraz V nad abecedou Σ je definován takto:

1. $\emptyset, \varepsilon, a$ jsou regulární výrazy pro všechna $a \in \Sigma$.
2. Jsou-li x, y regulární výrazy nad Σ, pak:
   - $(x + y)$ (sjednocení, alternativa),
   - $(x.y)$ (zřetězení),
   - $(x)^*$ (iterace)

   jsou regulární výrazy nad Σ.

Hodnota regulárního výrazu – Hodnota $h(x)$ regulárního výrazu x je definována takto:

1. $h(\emptyset) = \emptyset$ – prázdný jazyk,
2. $h(\varepsilon) = \{\varepsilon\}$ – neprázdný jazyk obsahující jeden prvek - ϵ,
3. $h(a) = \{a\}$,
4. $h(x + y) = h(x) \cup h(y)$,
5. $h(x.y) = h(x).h(y)$,
6. $h(x^*) = (h(x))^*$.

Hodnotou regulárního výrazu je (regulární) jazyk, který daný regulární výraz reprezentuje.

Axiomy pro regulární výrazy :

$A_1 : x + (y + z) = (x + y) + z$	(asociativnost sjednocení),
$A_2 : x + y = y + x$	(komutativnost sjednocení),
$A_3 : x + \emptyset = x$	($\emptyset$ je nulový prvek pro sjednocení),
$A_4 : x + x = x$	(idempotence sjednocení),
$A_5 : x.(y.z) = (x.y).z$	(asociativnost zřetězení),
$A_6 : \varepsilon x = x\varepsilon = x$	(ε je jednotkový prvek pro zřetězení),
$A_7 : \emptyset x = x\emptyset = \emptyset$	($\emptyset$ je nulový prvek pro zřetězení),
$A_8 : x.(y + z) = x.y + x.z$	(distributivnost zleva),
$A_9 : (x + y).z = x.z + y.z$	(distributivnost zprava),
$A_{10} : x^* = \varepsilon + x^*x$	
$A_{11} : x^* = (\varepsilon + x)^*$	
$A_{12} : x = x\alpha + \beta \Rightarrow x = \beta\alpha^*$	(řešení levé regulární rovnice),
$A_{13} : x = \alpha x + \beta \Rightarrow x = \alpha^*\beta$	(řešení pravé regulární rovnice).

Regulární rovnice – rozlišujeme dva druhy regulárních rovnic:

- **levá regulární rovnice** – ve tvaru:

  $$x = x\alpha + \beta$$

  část α nazýváme rekurzivní částí, zatímco část β. ukončovací. Řešení regulární rovnice lze snadno odvodit pomocí dosazovaní. Pro levou regulární rovnici je řešení ve tvaru:

  $$x = \beta\alpha^*$$

- **pravá regulární rovnice**

  $$x = \alpha x + \beta$$

  Pro pravou regulární rovnici je řešení ve tvaru:

  $$x = \alpha^*\beta$$

Soustava regulárních rovnic – obdobné jako soustava lineárních rovnic v matematice. Soustavu regulárních rovnic řešíme dosazovací metodou.

Derivace regulárních výrazů – Pro derivaci $\frac{d}{dx}$ regulárního výrazu V podle řetězce $x \in \Sigma^*$ platí:

$$\frac{dV}{dx} = V'$$
$$h(V') = \{y : xy \in h(V)\}$$

Derivaci si tedy představit jako odebírání předpony (dle které derivujeme) a to u všech řetězců, které reprezentuje daný regulární výraz (jenž jsou hodnotou daného regulárního výrazu).

Pravidla pro výpočet derivace regulárního výrazu V jsou následující:

1. $\frac{dV}{d\varepsilon} = V$
2. pro $a \in \Sigma$ platí:
   $\frac{d\varepsilon}{da} = \emptyset \quad \frac{d\emptyset}{da} = \emptyset$

   $$\frac{db}{da} = \begin{cases} \emptyset, \text{ jestliže } a \neq b \\ \varepsilon, \text{ jestliže } a = b \end{cases}$$

   $\frac{d(U+V)}{da} = \frac{dU}{da} + \frac{dV}{da}$

   $\frac{d(UV)}{da} = \frac{dU}{da}V + \{\frac{dV}{da} : \varepsilon \in h(U)\}$

   $\frac{d(V^*)}{da} = \frac{dV}{da}.V^*$

3. Pro $x = a_1a_2...a_n, a_i \in \Sigma$ platí

   $\frac{dV}{dx} = \frac{d}{da_n}(\frac{d}{da_{n-1}}(...\frac{d}{da_2}(\frac{dV}{da_1})...))$

Integrál regulárních výrazů – Integrál regulárního výrazu V podle řetězce $x \in \Sigma^*$ je definován takto:

$$h(\int V\ dx) = \{xy : y \in h(V)\}$$

Integrál si tedy představit jako přidávání předpony (dle které integrujeme) a to u všech řetězců, které reprezentuje daný regulární výraz (jenž jsou hodnotou daného regulárního výrazu).

Pravidla pro výpočet integrálu regulárního výrazu V jsou následující:

1. $\int V d\varepsilon = V$
2. pro $a \in \Sigma$ platí:
   $\int \varepsilon\ da = a$,
   $\int \emptyset\ da = \emptyset$,
   $\int b\ da = ab$,
   $\int (U + V)\ da = \int U\ da + \int V\ da$,
   $\int (U.V)\ da = aUV$,
   $\int V^*\ da = aV^*$.
3. pro $x = a_1 a_2 \cdots a_n \in \Sigma^*$ platí:
   $\int V\ dx = \int \cdots [\int (\int V\ da_n)\ da_{n-1}] \cdots da_1$.

Vztah mezi derivací a integrálem regulárního výrazu :

- Platí, že pokud daný regulární výraz V nejprve zintegrujeme dle x a poté zderivujeme dle x poté opět dostaneme původní regulární výraz V.

$$\frac{d}{dx} \int V\ dx = V$$

- Ne vždy ovšem platí, že pokud daný regulární výraz V nejprve zderivujeme dle x a poté zintegrujeme dle x poté opět dostaneme původní regulární výraz V.

$$\int \frac{dV}{dx}\ dx =^? V$$

  Neboť při derivaci můžeme nenávratně přijít o některé řetězce z jazyka, například:

  $V = (a(b + c) + ba,\ h(v) = \{ab, ac, ba\}$
  $\frac{dV}{da} = b + c$
  $\int \frac{dV}{da} da = a(b + c) \neq V$

  U integrálu proto zavádíme integrační konstantu Z:

$$\int V\ dx = xV + Z$$

$$\frac{dZ}{dx} = \emptyset$$

Úprava regulárních výrazů – na základě úprav rozlišujeme tři typy regulárních výrazů:

1. **Identické** – regulární výrazy x, y jsou identické, jestliže x a y jsou úplně stejné řetězce symbolů
2. **Ekvivalentní** – regulární výrazy x, y jsou ekvivalentní, jestliže mají stejnou hodnotu $h(x) = h(y)$, tj. reprezentují stejný regulární jazyk
3. **Podobné** – regulární výrazy x, y jsou podobné, jestliže se na sebe dají převést pomocí následujících identit:
   $x + x = x$
   $x + y = y + x$
   $(x + y) + z = x + (y + z)$
   $x + \emptyset = x$
   $x.\emptyset = \emptyset.x = \emptyset$
   $x.\varepsilon = \varepsilon.x = x$

Věty v axiomatické teorii regulárních výrazů :

$$
\begin{aligned}
V_1 &: \ \emptyset^* = \varepsilon \\
V_2 &: \ x^* + x = x^* \\
V_3 &: \ (x^*)^* = x^* \\
V_4 &: \ (x + y)^* = (x^* y^*)^* \\
V_5 &: \ x^* y = y + x^* xy \\
V_6 &: \ x^* y = y + xx^* y \\
V_7 &: \ x^* y = (x^n)^*.(y + xy + x^2 y + ... + x^{n-1} y) \\
V_8 &: \ \text{Jestliže } \varepsilon \in h(x), \text{ pak } xx^* = x^* \\
V_9 &: \ (xy)^* x = x(yx)^* \\
V_{10} &: \ (x + y)^* = (x^* + y^*)^*
\end{aligned}
$$