AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

Zbiory (języki) regularne

Teoria automatów i języków formalnych

Dr inż. Janusz Majewski
Katedra Informatyki

Zbiory (języki) regularne

Niech Σ będzie alfabetem.

Zbiór (język) regularny nad alfabetem Σ definiujemy następująco:

1) $\varnothing$ – zbiór pusty jest zbiorem regularnym,

2) $\{\varepsilon\}$ – zbiór zawierający łańcuch pusty jest zbiorem regularnym,

3) $\{a\}$ – ($\forall a \in \Sigma$) zbiór zawierający łańcuch złożony z pojedynczego symbolu alfabetu jest zbiorem regularnym,

4) jeśli P i Q są zbiorami regularnymi nad Σ to zbiorami regularnymi są także:

   a) $P \cup Q$ – suma teoriomnogościowa zbiorów P i Q,

   b) PQ – złożenie (konkatenacja) zbiorów P i Q,

   c) P^* - domknięcie Kleene'ego zbioru P.

5) nic innego poza tym, co wynika z punktów (1) – (4), nie jest zbiorem regularnym.

Wyrażenia regularne (1)

Wyrażenia regularne służą do uproszczonego oznaczania zbiorów regularnych. Niech Σ będzie alfabetem. Wyrażenia regularne nad alfabetem Σ definiujemy następująco:

1) $\varnothing$ – jest wyrażeniem regularnym oznaczającym zbiór pusty $\varnothing$ będący zbiorem regularnym,

2) $\boldsymbol{\varepsilon}$ – jest wyrażeniem regularnym oznaczającym zbiór zawierający łańcuch pusty $\{\varepsilon\}$ będący zbiorem regularnym,

3) ***a*** – ($\forall a \in \Sigma$) jest wyrażeniem regularnym oznaczającym zbiór zawierający łańcuch złożony z pojedynczego symbolu alfabetu będący zbiorem regularnym,

4) jeśli ***p*** i ***q*** są wyrażeniami regularnymi oznaczającymi odpowiednio zbiory regularne P i Q nad Σ to wyrażeniami regularnymi są także:

- ***a)*** ***p|q*** – wyrażenie regularne oznaczające $P \cup Q$ – sumę teoriomnogościową zbiorów P i Q będącą zbiorem regularnym,
- ***b)*** ***pq*** – wyrażenie regularne oznaczające PQ – złożenie (konkatenację) zbiorów P i Q będące zbiorem regularnym,
- ***c)*** ***p**** - wyrażenie regularne oznaczające P^* - domknięcie Kleene'ego zbioru P będące zbiorem regularnym.

5) nic innego poza tym, co wynika z punktów (1) – (4), nie jest wyrażeniem regularnym.

Przykłady

Przykład:

Niech Σ = {a, b}. Zbiorem regularnym nad Σ jest np. zbiór:

$\{\varepsilon, a, ab, abb, abbb, abbbb, ...\} = \{\varepsilon\} \cup \{a\}\{b\}^*$

Ten zbiór regularny zapisujemy w formie wyrażenia regularnego jako:

ε|ab*.

Przykład:

Wyrażenie regularne:

(0|1)*011

odpowiada zbiorowi regularnemu:

$(\{0\} \cup \{1\})^* \{0\}\{1\}\{1\} = \{0, 1\}^* \{011\}$

będącemu dowolnym ciągiem zer i jedynek zakończonym sekwencją: 011.

Wyrażenia regularne (2)

- Dwa wyrażenia **p i q** regularne są równe (równoważne), gdy odpowiadające im zbiory regularne *P* i *Q* są równe (identyczne).
- W zapisie wyrażeń regularnych można stosować nawiasy.
- Zapisując wyrażenia regularne stosujemy następujące priorytety operatorów:
  - ( ) - najwyższy,
  - *
  - ▪ - (konkatenacja)
  - | - najniższy.

Tożsamości

Niech **p**, **q** i **r** będą dowolnymi wyrażeniami regularnymi. Prawdziwe są następujące zależności i tożsamości:

$$p \mid q = q \mid p$$
$$p \mid (q \mid r) = (p \mid q) \mid r$$
$$p(qr) = (pq)r$$
$$pq \mid pr = p(q \mid r)$$
$$pq \mid rq = (p \mid r)q$$
$$\varepsilon p = p\varepsilon = p$$
$$\varnothing p = p\varnothing = \varnothing$$
$$\varepsilon^* = \varepsilon$$
$$\varnothing^* = \varepsilon$$

$$p^* = p \mid p^* = (p \mid \varepsilon)^*$$
$$(p^*)^* = p^{**} = p^*$$
$$p \mid p = p$$
$$p \mid \varnothing = p$$
$$\varepsilon \mid p^* = p^*$$
$$\varepsilon \mid pp^* = p^*$$
$$\varepsilon \mid p^*p = p^*$$
$$pqq^* \mid pq^* = pq^*$$
$$(p \mid q)^* = (p^* \mid q^*)^* = (p^*q^*)^*$$

Przykłady (1)

Przykład:

abb*|ab* = ab*

bo:

abb*|ab* = {ab}{b}* ∪ {a}{b}* = {ab, abb, ...} ∪ {a, ab, abb, ...} =
= {a, ab, abb, ...} = {a}{b}*

Przykład:

(ab|a)*a = a(ba|a)*

bo:

(ab|a)* - to łańcuchy zbudowane z liter a i b, rozpoczynające się literą a, w których żadne dwie litery b, o ile w ogóle występują, nie występują obok siebie,

(ba|a)* - to łańcuchy zbudowane z liter a i b, kończące się literą a, w których żadne dwie litery b, o ile w ogóle występują, nie występują obok siebie,

(ab|a)*a oraz **a(ba|a)*** - to łańcuchy zbudowane z liter a i b, rozpoczynające się i kończące się literą a, w których żadne dwie litery b, o ile w ogóle wystepują, nie występują obok siebie.

Przykłady (2)

Przykład:

(a|b)* ≠ a*|b*

bo:

(a|b)* = {a, b}* = {ε, a, b, aa, ab, ba, bb, aaa, aab, ...}

a*|b* = {a}* ∪ {b}* = {ε, a, aa, aaa, ...} ∪ {ε, b, bb, bbb, ...} =
={ε, a, b, aa, bb, aaa, bbb, ...}

Przykład:

b(ab|b)* ≠ aa*b(aa*b)*

bo:

b(ab|b)* - to łańcuchy zbudowane z liter a i b, rozpoczynające się i kończące się literą b, w których żadne dwie litery a, nie występują obok siebie,

aa*b(aa*b)* - to łańcuchy zbudowane z liter a i b, rozpoczynające się literą a, kończące się literą b, w których żadne dwie litery b nie występują obok siebie.