

Jméno:

UČO:

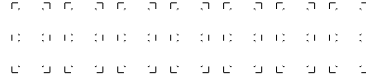
Skupina:



líst



učo



body



Oblast strojově snímaných informací. Svě učo a číslo lístu vyplňte zleva dle vzoru číslic. Jinak do této oblasti nezasahujte.

0 1 2 3 4 5 6 7 8 9

1. [2 body]

Nechť L je regulární jazyk nad libovolnou abecedou Σ a unární operace $removeOne(L)$ je definována následovně:

$$removeOne(L) = \{uv \in \Sigma^* \mid \exists x \in \Sigma \text{ takové, že } u xv \in L\}$$

Intuitivně je tedy $removeOne(L)$ jazyk obsahující všechna slova w taková, že w vznikne odstraněním jednoho písmene z nějakého slova délky alespoň jedna z jazyka L . Například:

$$\begin{aligned} removeOne(\{b, abab\}) &= \{\varepsilon, bab, aab, abb, aba\} \\ removeOne(\{\varepsilon, ab, bba, ababab\}) &= \{a, b, ba, bb, babab, aabab, abbab, abaab, ababb, ababa\} \\ removeOne(\{a\}^+) &= \{a\}^* \\ removeOne(\{aa\}^*) &= \{a\} \cdot \{aa\}^* \end{aligned}$$

Vášim úkolem je rozhodnout, zda je jazyk $removeOne(L)$ regulární, tedy že třída regulárních jazyků je uzavřená na operaci $removeOne$. Vaši odpověď dokažte, a to tak, že:

- Pokud rozhodnete, že není, najděte regulární jazyk L takový, že jazyk $removeOne(L)$ regulární není.
- Pokud rozhodnete, že je, dokažte tvrzení například s pomocí známých uzávěrových vlastností třídy regulárních jazyků prezentovaných na přednášce, nebo konstruktivně popsáním algoritmu na transformaci nějakého formalizmu pro popis regulárních jazyků.

Třída regulárních jazyků je uzavřená na operaci $removeOne$. To můžeme dokázat popisem algoritmu na konstrukci konečného automatu \mathcal{A} rozpoznávajícího jazyk $removeOne(L)$.

Základní myšlenka konstrukce

Budeme konstruovat konečný automat pro výsledný jazyk. Vyjdeme z DFA pro jazyk L (ten existuje, protože L je regulární) a jeho kopie. Kdykoli načteme nějaký prefix u slova $w = u xv$ z jazyka L , budeme moci namísto přechodu pod znakem x (ze stavu q do r) provést přechod pod ε (tím dojde k odstranění znaku x) z q do stavu odpovídajícímu r , ale v kopii původního automatu, a pak dokončit načítání příslušného sufixu v v kopii.

Myšlenka konstrukce konečného automatu – podrobněji

Mějme deterministický konečný automat \mathcal{A}_L pro jazyk L a jeho kopii \mathcal{A}'_L (stavy automatu \mathcal{A}'_L mají stejné označení, jako odpovídající stavy v \mathcal{A}_L , ale navíc jsou označeny $'$, tedy například q je stav automatu \mathcal{A}_L , q' je odpovídající stav \mathcal{A}'_L).

Nyní spojíme uvedené automaty do jednoho automatu \mathcal{A} . Pro každý přechod ze stavu q do stavu r automatu \mathcal{A}_L vedeme přechod pod ε ze stavu q automatu \mathcal{A}_L do stavu r' automatu \mathcal{A}'_L . Stav r' odpovídá stavu r z automatu \mathcal{A}_L .

Tímto jsme spojili automaty a dále je potřeba zvolit počáteční stav a akceptující stavy \mathcal{A} . Jako počáteční stav \mathcal{A} vybereme stav, který byl počátečním stavem automatu \mathcal{A}_L . Koncové stavy budou všechny koncové stavy automatu \mathcal{A}'_L , tím vynutíme vynechání alespoň jednoho písmene.

Jméno:

UČO:

Skupina:

0007

list

2

učo

body

Oblast strojově snímaných informací. Svě učo a číslo listu vyplňte zleva dle vzoru číslic. Jinak do této oblasti nezasahujte.

0123456789

Mohlo by se zdát, že automat \mathcal{A}'_L není potřeba a přechody pod ε by stačilo vytvořit jednoduše v automatu \mathcal{A}_L . Tím bychom však vytvořili jazyk, který dovoluje odstanit více znaků v každém slově z jazyka L .

Formální zápis konstrukce konečného automatu

DFA pro jazyk L je dán pěticí $\mathcal{A}_L = (Q_L, \Sigma, \delta_L, q_{L0}, F_L)$. Pro jazyk $removeOne(L)$ vytvoříme NFA s ε -kroky $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$, kde

- $Q = (\{1, 2\} \times Q_L)$,
- $q_0 = (1, q_{L0})$,
- $F = \{2\} \times F_L$,
- $\delta((1, q), \varepsilon) = \{(2, \delta_L(q, x)) \mid x \in \Sigma\}$, pro všechna $q \in Q_L$,
- $\delta((2, q), \varepsilon) = \emptyset$, pro všechna $q \in Q_L$,
- $\delta((i, q), a) = \{(i, \delta_L(q, a))\}$, pro všechna $i \in \{1, 2\}$, $a \in \Sigma$ a $q \in Q_L$.

Poznámka: Stavů automatu \mathcal{A} jsou označeny uspořádanými dvojicemi, kde první složka určuje, ve které části slova $u \cdot x \cdot v$ ($uxv \in L$) se nacházíme. Dvojice tvaru $(1, q)$ označují stavy pro část u , dvojice $(2, q)$ pro část v . K odstranění znaku x dojde provedením ε -kroku.