Matematička teorija računarstva
Zadaća 2

Filip Nikšić
fniksic@gmail.com

10. ožujka 2008.

Originalni dio

Zadatak 1. Opišite riječima najmanje tri jezika te za njih konstruirajte regularne izraze.

Rješenje. Neka je $\Sigma = \{0, 1, \ldots, 9, a, b, \ldots, z\}$ skup koji sadrži znamenke i mala slova engleske abecede.

- $L_1 = \{w \in \Sigma^\star \mid w \text{ sadrži riječ } \texttt{filip} \text{ kao podriječ}\}$

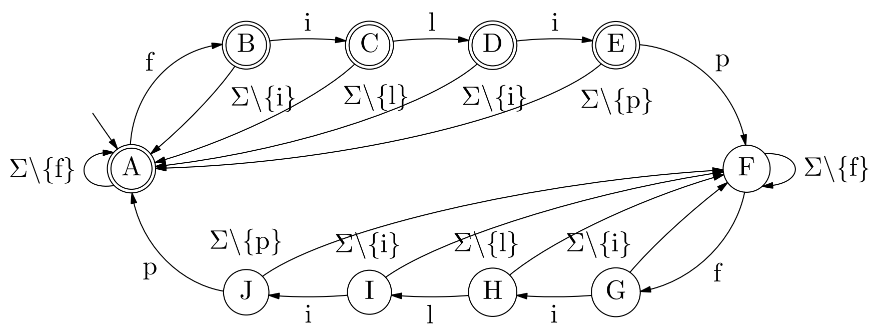
  Regularan izraz za ovaj jezik je jednostavan: $\Sigma^\star\texttt{filip}\Sigma^\star$.

- $L_2 = \{w \in \{0, 1\}^\star \mid \text{iza svake } 0 \text{ nalazi se točno jedna } 1\}$

  Regularan izraz je opet jednostavan: $1^\star(01)^\star$.

- $L_3 = \{w \in \Sigma^\star \mid w \text{ sadrži paran broj pojava podriječi } \texttt{filip}\}$

  Ovaj put regularan izraz nije toliko očit. No nije toliko teško konstruirati DKA koji prepoznaje L_3.



Konstruiramo skupovne jednadžbe ($\Sigma\backslash x$ je skraćena oznaka za $\Sigma\backslash\{x\}$):

$$
\begin{aligned}
A &= (\Sigma\backslash\mathtt{f})A + \mathtt{f}B + \varepsilon & F &= (\Sigma\backslash\mathtt{f})F + \mathtt{f}G \\
B &= (\Sigma\backslash\mathtt{i})A + \mathtt{i}C + \varepsilon & G &= (\Sigma\backslash\mathtt{i})F + \mathtt{i}H \\
C &= (\Sigma\backslash\mathtt{l})A + \mathtt{l}D + \varepsilon & H &= (\Sigma\backslash\mathtt{l})F + \mathtt{l}I \\
D &= (\Sigma\backslash\mathtt{i})A + \mathtt{i}E + \varepsilon & I &= (\Sigma\backslash\mathtt{i})F + \mathtt{i}J \\
E &= (\Sigma\backslash\mathtt{p})A + \mathtt{p}F + \varepsilon & J &= (\Sigma\backslash\mathtt{p})F + \mathtt{p}A
\end{aligned}
$$

Uvrštavanjem unatrag J u I, pa I u H, pa H u G, pa G u F i korištenjem Ardenove leme otkrivamo da je

$$F = \big(\Sigma\backslash\mathtt{f} + \mathtt{f}(\Sigma\backslash\mathtt{i}) + \mathtt{fi}(\Sigma\backslash\mathtt{l}) + \mathtt{fil}(\Sigma\backslash\mathtt{i}) + \mathtt{fili}(\Sigma\backslash\mathtt{p})\big)^\star \mathtt{filip}A$$

Nastavljamo uvrštavati unatrag: F u E, E u D, D u C, C u B i konačno B u A.

$$
\begin{aligned}
A = {} & \big(\Sigma\backslash\mathtt{f} + \mathtt{f}(\Sigma\backslash\mathtt{i}) + \mathtt{fi}(\Sigma\backslash\mathtt{l}) + \mathtt{fil}(\Sigma\backslash\mathtt{i}) + \mathtt{fili}(\Sigma\backslash\mathtt{p})\big)A \\
& + \mathtt{filip}\big(\Sigma\backslash\mathtt{f} + \mathtt{f}(\Sigma\backslash\mathtt{i}) + \mathtt{fi}(\Sigma\backslash\mathtt{l}) + \mathtt{fil}(\Sigma\backslash\mathtt{i}) + \mathtt{fili}(\Sigma\backslash\mathtt{p})\big)^\star \mathtt{filip}A \\
& + \mathtt{fili} + \mathtt{fil} + \mathtt{fi} + \mathtt{f} + \varepsilon
\end{aligned}
$$

Još jednim iskorištavanjem Ardenove leme dobivamo konačan regularan izraz. Zbog kraćeg zapisa uvodimo oznaku $\Lambda = \big(\Sigma\backslash\mathtt{f} + \mathtt{f}(\Sigma\backslash\mathtt{i}) + \mathtt{fi}(\Sigma\backslash\mathtt{l}) + \mathtt{fil}(\Sigma\backslash\mathtt{i}) + \mathtt{fili}(\Sigma\backslash\mathtt{p})\big)$.

$$A = \big(\Lambda + \mathtt{filip}\,\Lambda^\star\,\mathtt{filip}\big)^\star(\mathtt{fili} + \mathtt{fil} + \mathtt{fi} + \mathtt{f} + \varepsilon)$$

■

Zadatak 2 (10 bodova)**.** Riješite sustav skupovnih jednadžbi:

$$
\begin{aligned}
U &= \varepsilon + \mathtt{d}V + \mathtt{a}\Lambda + \mathtt{g}U \\
V &= \mathtt{f}U + \mathtt{e}V \\
\Lambda &= \mathtt{c}U + \mathtt{b}\Lambda
\end{aligned}
$$

Rješenje. Prvo izražavamo V i Λ preko U i uvrštavamo to u jednadžbu za U:

$$
\begin{aligned}
V &= \mathtt{e}^\star\mathtt{f}U \\
\Lambda &= \mathtt{b}^\star\mathtt{c}U \\
U &= \varepsilon + \mathtt{de}^\star\mathtt{f}U + \mathtt{ab}^\star\mathtt{c}U + \mathtt{g}U
\end{aligned}
$$

Korištenjem Ardenove leme dobivamo U pa to koristimo da dobijemo V i Λ:

$$
\begin{aligned}
U &= (\mathtt{de}^\star\mathtt{f} + \mathtt{ab}^\star\mathtt{c} + \mathtt{g})^\star \\
V &= \mathtt{e}^\star\mathtt{f}(\mathtt{de}^\star\mathtt{f} + \mathtt{ab}^\star\mathtt{c} + \mathtt{g})^\star \\
\Lambda &= \mathtt{b}^\star\mathtt{c}(\mathtt{de}^\star\mathtt{f} + \mathtt{ab}^\star\mathtt{c} + \mathtt{g})^\star
\end{aligned}
$$

■

Zadatak 3 (20 bodova). Neka su Σ i Γ dvije (ne nužno različite) abecede i neka je $f\colon \Sigma \to \Gamma^\star$ proizvoljna funkcija. Proširimo f na riječi nad abecedom Σ:

$$f(\varepsilon) := \varepsilon, \quad f(w_1 w_2 \dots w_n) := f(w_1)f(w_2)\dots f(w_n), \quad w_1, \dots, w_n \in \Sigma$$

Dokažite ili opovrgnite: Za regularan jezik $L \subseteq \Sigma^\star$ slika $f(L) := \{f(v) \mid v \in L\}$ je također regularan jezik.

Rješenje. Definirajmo transformaciju regularnog izraza R nad Σ u regularan izraz R' nad Γ na sljedeći način: za sve $x \in \Sigma$ svaku pojavu znaka x u R zamijenimo sa $f(x)$.

Indukcijom po duljini izraza R dokazujemo da je $L(R') = f(L(R))$. Za $R = \varnothing$ i $R = \varepsilon$ tvrdnja očito vrijedi. Za $R = x$ gdje je $x \in \Sigma$ je $R' = f(x)$ pa zbog $f(L(x)) = f(\{x\}) = \{f(x)\} = L(f(x))$ tvrdnja vrijedi.

Pretpostavimo da tvrdnja vrijedi za sve regularne izraze čija je duljina manja od n za neki $n > 1$. Neka je R izraz duljine n. Tad je R oblika $R_1 + R_2$, R_1R_2 ili $R_1^\star$.

U prvom slučaju R' je očito oblika $R_1' + R_2'$. Vrijedi

$$f(L(R)) = f(L(R_1 + R_2)) = f(L(R_1) \cup L(R_2)) = f(L(R_1)) \cup f(L(R_2)).$$

Duljina R_1 i R_2 strogo je manja od n pa po pretpostavci indukcije dalje vrijedi

$$f(L(R)) = L(R_1') \cup L(R_2') = L(R_1' + R_2') = L(R').$$

U slučaju kad je R oblika R_1R_2 očito je R' oblika $R_1'R_2'$. Imamo:

$$f(L(R)) = f(L(R_1R_2)) = f(L(R_1)L(R_2))$$

Zbog definicije funkcije f vrijedi $f(uv) = f(u)f(v)$ gdje su u i v riječi iz $\Sigma^\star$. To povlači $f(L(R_1)L(R_2)) = f(L(R_1))f(L(R_2))$ pa uvažavajući i pretpostavku indukcije vrijedi:

$$f(L(R)) = L(R_1')L(R_2') = L(R_1'R_2') = L(R').$$

Treći slučaj je zbog definicije zvijezde kombinacija prethodna dva pa se dokazuje analogno.

Tvrdnja zadatka sad direktno slijedi iz dokazanog. Naime, za regularan jezik L koji je opisan regularnim izrazom R, slika $f(L)$ je opisana regularnim izrazom R' pa je ona također regularan jezik. ∎

Zadatak 4 (15 bodova). Dokaži ili opovrgni: regularni izrazi $R = a^\star(ba + aa)^\star$, $S = a^\star\big(ba(aa)^\star\big)^\star$ i $T = a^\star\big(ba(aa + ba)^\star + \varepsilon\big)$ svi opisuju isti regularan jezik.

Rješenje. Dokažimo da R i S opisuju iste jezike. Naime, neka je w riječ iz jezika opisanog izrazom $R = a^\star(aa + ba)^\star$. Ukoliko w ne sadrži znak b, očito je $w \in a^\star \subseteq a^\star\big(ba(aa)^\star\big)^\star$. Ako w sadrži b, onda je w oblika

$$w = a^k(aa)^{l_1}(ba)^{m_1} \dots (aa)^{l_n}(ba)^{m_n}(aa)^{l_{n+1}}$$

pri čemu je BSO $m_i \neq 0$. (Ako je $m_i = 0$, spojimo $(\mathsf{aa})^{l_i}(\mathsf{aa})^{l_{i+1}} = (\mathsf{aa})^{l_i + l_{i+1}}$.) Sad imamo:

$$\begin{aligned} w &= \mathsf{a}^{k+2l_1}(\mathsf{ba})^{m_1-1}\big(\mathsf{ba}(\mathsf{aa})^{l_2}\big)\dots(\mathsf{ba})^{m_n-1}\big(\mathsf{ba}(\mathsf{aa})^{l_{n+1}}\big) \\ &\in \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{aa})^\star\big)^{m_1-1}\big(\mathsf{ba}(\mathsf{aa})^\star\big)\dots\big(\mathsf{ba}(\mathsf{aa})^\star\big)^{m_n-1}\big(\mathsf{ba}(\mathsf{aa})^\star\big) \\ &= \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{aa})^\star\big)^{m_1+\dots+m_n} \\ &\subseteq \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star \end{aligned}$$

Obrnuto je lakše. Naime

$$\mathsf{aa} \subseteq \mathsf{ba} + \mathsf{aa} \qquad \Longrightarrow \qquad (\mathsf{aa})^\star \subseteq (\mathsf{ba} + \mathsf{aa})^\star .$$

Nadalje, $\mathsf{ba} \subseteq \mathsf{ba} + \mathsf{aa}$ pa imamo

$$\begin{aligned} \mathsf{ba}(\mathsf{aa})^\star &\subseteq (\mathsf{ba}+\mathsf{aa})(\mathsf{ba}+\mathsf{aa})^\star \subseteq (\mathsf{ba}+\mathsf{aa})^\star \\ \big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star &\subseteq \big((\mathsf{ba}+\mathsf{aa})^\star\big)^\star = (\mathsf{ba}+\mathsf{aa})^\star \\ \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star &\subseteq \mathsf{a}^\star(\mathsf{ba}+\mathsf{aa})^\star \end{aligned}$$

Dokažimo da S i T opisuju iste jezike. Prvo pokažimo da je $\big(\mathsf{ba}(\mathsf{aa})^\star\big)^k \subseteq \mathsf{ba}(\mathsf{ba}+\mathsf{aa})^\star$ za svaki $k > 0$. Naime, za $k = 1$ tvrdnja očito vrijedi. Pretpostavimo da vrijedi za neki $k > 0$. Tad je

$$\begin{aligned} \big(\mathsf{ba}(\mathsf{aa})^\star\big)^{k+1} &= \big(\mathsf{ba}(\mathsf{aa})^\star\big)^k\mathsf{ba}(\mathsf{aa})^\star \\ &\subseteq \mathsf{ba}(\mathsf{ba}+\mathsf{aa})^\star\mathsf{ba}(\mathsf{aa})^\star \\ &\subseteq \mathsf{ba}(\mathsf{ba}+\mathsf{aa})^\star(\mathsf{ba}+\mathsf{aa})(\mathsf{ba}+\mathsf{aa})^\star \\ &\subseteq \mathsf{ba}(\mathsf{ba}+\mathsf{aa})^\star \end{aligned}$$

Iz ovog zaključujemo da je

$$\begin{aligned} \big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star &\subseteq \mathsf{ba}(\mathsf{ba}+\mathsf{aa})^\star + \varepsilon \\ \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star &\subseteq \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{ba}+\mathsf{aa})^\star + \varepsilon\big) \end{aligned}$$

U obratnom slučaju pokažimo da je $\mathsf{ba}(\mathsf{aa}+\mathsf{ba})^\star \subseteq \big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star$. Naime, riječ w iz jezika s lijeve strane je oblika

$$w = \mathsf{ba}(\mathsf{aa})^{k_1}(\mathsf{ba})^{l_1}\dots(\mathsf{aa})^{k_n}(\mathsf{ba})^{l_n}(\mathsf{aa})^{k_{n+1}}$$

s tim da kao i u prvom dijelu zadatka možemo BSO pretpostaviti $l_i \neq 0$. Daljnje zaključivanje je također sasvim analogno onom u prvom dijelu zadatka. Imamo $w \in \big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star$. Kad sve skupimo dobijemo:

$$\begin{aligned} \mathsf{ba}(\mathsf{aa}+\mathsf{ba})^\star + \varepsilon &\subseteq \big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star \\ \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{aa}+\mathsf{ba})^\star + \varepsilon\big) &\subseteq \mathsf{a}^\star\big(\mathsf{ba}(\mathsf{aa})^\star\big)^\star \end{aligned}$$

■