

Formele Talen - Inleveropgaven I

Martijn Vermaat
mvermaat@cs.vu.nl

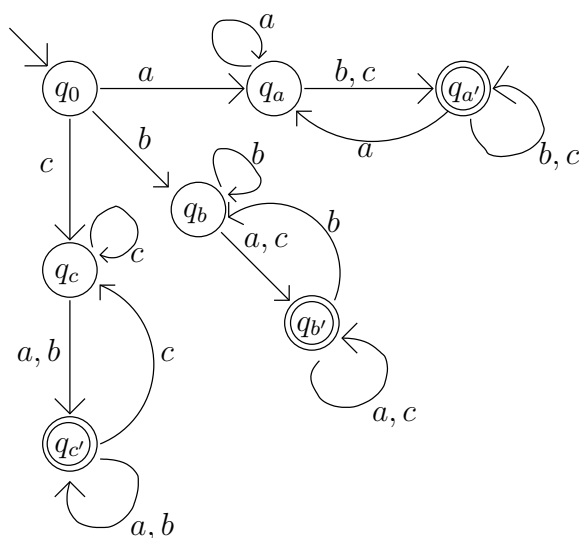
16 november 2004

Opgave 1

a) Een grammatica die L genereert:

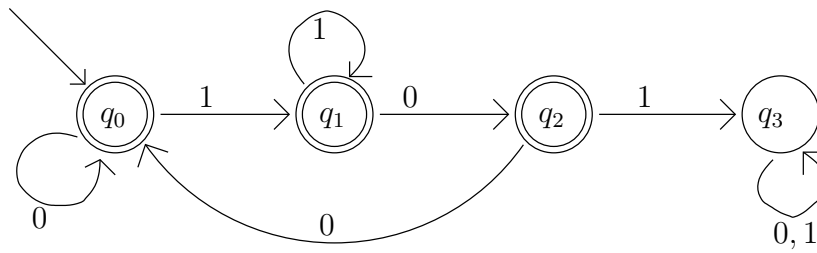
$$\begin{aligned} S &\rightarrow aAb \mid aAc \mid bAa \mid bAc \mid cAa \mid cAb \\ A &\rightarrow \lambda \mid aA \mid bA \mid cA \end{aligned}$$

b) Een dfa die L accepteert:

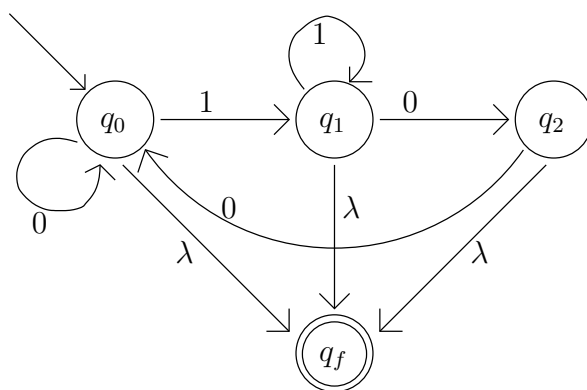


Opgave 2

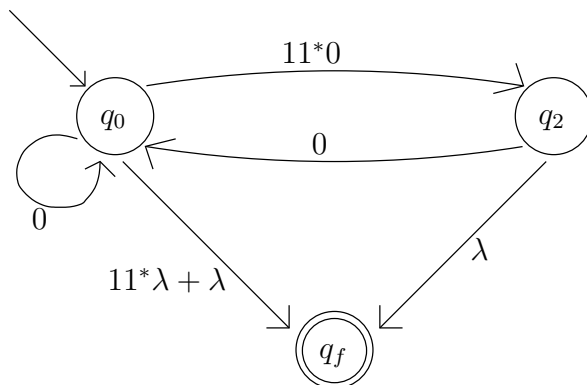
a) Een dfa die L accepteert:



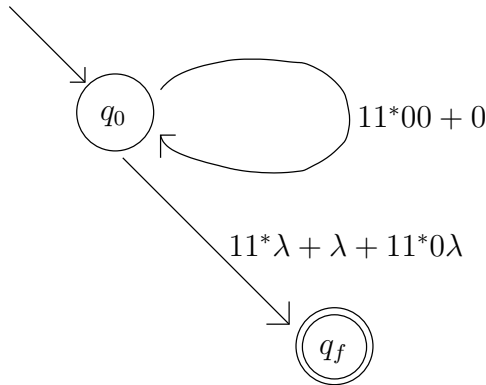
- b) We maken er eerst een nfa van waarbij we de trapstate q_3 weg laten en de drie accepterende states vervangen door één accepterende state:



Vervolgens halen we q_1 weg volgens het algoritme op pagina's 82 en 83 van het boek:



En zo ook q_2 :



Volgens het recept uit datzelfde algoritme kunnen we uit deze nfa de volgende reguliere expressie construeren:

$$r = (11^*00 + 0)^* (11^*\lambda + \lambda + 11^*0\lambda)$$

Opgave 3

We bewijzen dat de taal

$$\{a^n \mid n \text{ is geen priemgetal}\}$$

niet regulier is door met behulp van de pompstelling te bewijzen dat het complement van deze taal niet regulier is.

Bewijs. We nemen aan dat de taal

$$L = \{a^n \mid n \text{ is een priemgetal}\}$$

regulier is. Omdat L een oneindige taal is (er zijn oneindig veel priemgetallen) kunnen we de pompstelling gebruiken. Volgens deze stelling bestaat er een geheel getal m , zó, dat voor iedere $w \in L$ met $|w| \geq m$ er een opdeling

$$w = xyz$$

bestaat met $|xy| \leq m$ en $|y| \geq 1$ waarbij

$$w_i = xy^i z$$

ook in L zit (voor alle $i \geq 0$).

Gegeven het gehele getal m , bekijken we de string

$$w = a^n \text{ (} n \text{ is het } m^{\text{de}} \text{ priemgetal).}$$

Nu bekijken we met behulp van de genoemde opdeling xyz van w de string

$$w_n = xy^n z \text{ (} n = |w| + 1 \text{).}$$

Nu hebben we $|w_n| = |w| + (|w| * |n|)$, waarbij we zeker weten dat deze lengte deelbaar is door $|w|$ (met als resultaat $|n| + 1$). En dus zit w_n niet in L . Maar de pompstelling zegt dat w_n wel in L zit. Dit is een tegenspraak.

Uit deze tegenspraak concluderen we dat onze aanname niet juist was en dus dat L regulier is.

Doordat de eigenschap ‘regulier’ gesloten is onder de operatie complement¹, is ook het complement

$$\{a^n \mid n \text{ is geen priemgetal}\}$$

van L niet regulier. Dit wilden we bewijzen. ■

Opgave 4

Omdat L_1 een reguliere taal over $\{a, b\}$ is, bestaat er een reguliere expressie r_1 die precies L_1 beschrijft.

De reguliere expressie r_2 definiëren we door herschrijfgeregels op r_1 toe te passen:

$$r_2 \mapsto r'_1 \tag{1}$$

$$(\psi\omega)' \mapsto \psi'\omega + \psi\omega' \tag{2}$$

$$(\psi + \omega)' \mapsto \psi' + \omega' \tag{3}$$

$$(\psi^*)' \mapsto \psi^*\psi'\psi^* \tag{4}$$

$$\alpha' \mapsto \lambda \tag{5}$$

$$\lambda' \mapsto \emptyset \tag{6}$$

$$\emptyset' \mapsto \emptyset \tag{7}$$

Waarbij ψ en ω reguliere expressies zijn en α een element uit $\{a, b\}$. De taal die door r_2 beschreven wordt is regulier.

We zullen nu bewijzen dat de reguliere expressie r_2 precies L_2 beschrijft.

Bewijs. We bewijzen met inductie naar de structuur van r_1 . Als basisstap nemen we r_1 gelijk aan een element uit $\{a, b, \lambda, \emptyset\}$. Volgens de definitie van L_2 (uit de opgave) is $L_2 = \{\lambda\}$ indien $r_1 \in \{a, b\}$ en $L_2 = \emptyset$ als $r_1 \in \{\lambda, \emptyset\}$.

Regel 5 uit onze herschrijfgeregels zegt dat r_2 geschreven kan worden als λ wanneer $r_1 \in \{a, b\}$. Regels 6 en 7 zeggen dat r_2 geschreven kan worden als \emptyset wanneer $r_1 \in \{\lambda, \emptyset\}$. Deze twee beschrijven precies de taal $\{\lambda\}$ respectievelijk \emptyset en dus hebben we hiermee ons basisgeval bewezen.

¹Eigenlijk gebruiken we hier dat het ‘niet-regulier’ zijn gesloten is onder de complement operatie. Dit is ook algemeen geldig, omdat een dubbel complement de originele taal geeft.

Onze inductiestap bestaat uit het bewijzen dat we met onze herschrijfgeregels voor iedere reguliere expressie r_1 een reguliere expressie r_2 kunnen maken die precies L_2 beschrijft waarbij we aan mogen nemen dat dit al bewezen is voor iedere subexpressie van r_1 . We behandelen alle gevallen van de structuur van r_1 apart (met uitzondering van de gevallen waarin deze precies a , b , λ , of \emptyset is, want dat hebben we in de basisstap gedaan).

Concatenatie Wanneer r_1 van de vorm $\psi\omega$ is (met ψ en ω reguliere expressies), dan zien we aan de definitie van L_2 dat we L_2 krijgen door in de strings van L_1 een letter weg te laten uit *ofwel* het deel dat beschreven wordt door ψ *ofwel* het deel dat beschreven wordt door ω . Het is niet moeilijk in te zien dat dat precies is wat regel 2 beschrijft.

Keuze Evenzo, als r_1 van de vorm $\psi+\omega$ is, zegt de definitie van L_2 dat deze die strings bevat die beschreven worden door *ofwel* ψ met daaruit een letter weggelaten, *ofwel* ω met daaruit een letter weggelaten. Regel 3 beschrijft precies dit.

Ster-afsluiting Als laatste geval bekijken we r_1 wanneer deze van de vorm ψ^* is. Het is al laat, dus ik laat het erbij dat het niet moeilijk is om in te zien dat L_2 precies die strings bevat die bestaan uit het nul of meer keer herhalen van ψ , gevolgd door ψ met daaruit een letter weggelaten, gevolgd door wederom het nul of meer keer herhalen van ψ . Deze strings worden precies beschreven door regel 4.

Nu we hebben bewezen dat de reguliere expressie r_2 precies de taal L_2 beschrijft, volgt onmiddellijk dat L_2 een reguliere taal is.

■