Logische Complexiteit
Deeltoets

Jeroen Goudsmit

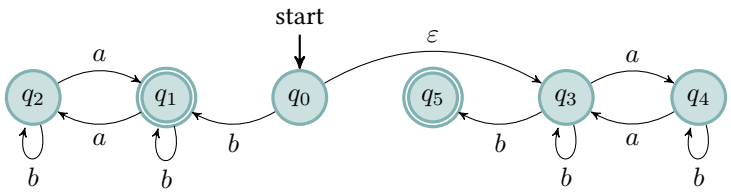
donderdag 1 maart 2012
11:00 – 13:00

Hieronder de uitwerkingen van de deeltoets. Er zijn hier meer bewijzen gegeven in iets meer detail dan er van je verlangt werd. Dit document is bedoelt je te laten zien hoe je iedere vraag *had kunnen* beantwoorden, niet hoe je per se had moeten antwoorden. Als je fouten ziet, meldt dat dan a.u.b. bij Jeroen Goudsmit.

Opgave 1 — Reguliere taal herkennen

De hieronder gegeven NFA herkent de onderstaande taal over het alfabet $\Sigma := \{a, b\}$.

$$\mathcal{L} = \{\ w \in \Sigma^* \mid w \text{ bevat een even aantal } a\text{'s en begint óf eindigt met een } b\ \} \tag{1}$$



Het idee is dat een de automaat in het begin "kiest" of het woord begint met een b. Zo ja, dan sprint 'ie naar links, en accepteert als er een even aantal a's voorbij gekomen is. Zo nee, dan kun je na een even aantal a's via een b springen naar een accepterende toestand waar verder niks mogelijk is. Dit pas leidt dus tot acceptatie precies wanneer je een even aantal a's bevat en eindigt op een b.

Opgave 2 — Context-vrije talen herkennen

Geef CFG's voor de volgende twee talen over het alfabet $\Sigma := \{d, e, l, o, s, t\}$. Bij beide talen dien je te beargumenteren waarom de door jou gegeven grammatica klopt. Bij één van de twee moet je ook bewijzen dat de grammatica klopt, je mag zelf kiezen welke taal dat wordt.

(i) $\{\ (deel)^n (toets)^n \mid n \in \mathbb{N}\ \}$

(ii) $\{\ e^n o^m \mid n, m \in \mathbb{N} \text{ en } n \geq m\ \}$.

Lemma 1. *De taal* $\mathcal{L} = \{\ (deel)^n(toets)^n \mid n \in \mathbb{N}\ \}$ *wordt herkent door de* CFG G *als gegeven hieronder.*

$$\mathsf{S} \to deel\ \mathsf{S}\ toets \mid \varepsilon$$

Bewijs. We willen bewijzen dat als $w \in \Sigma^*$, dan geldt $w \in \mathcal{L}(G)$ dan en slechts dan als $w \in \mathcal{L}$.

Laten we eerst de implicatie van rechts naar links bewijzen. We nemen aan dat $w \in \mathcal{L}$. Er volgt dat er een $n \in \mathbb{N}$ is zodanig dat $w = (deel)^n(toets)^n$. We bewijzen met inductie naar n dat $\mathsf{S} \overset{*}{\Rightarrow} w$. In het geval dat $n = 0$ dan geldt $w = \varepsilon$, en het is waar dat $\mathsf{S} \overset{*}{\Rightarrow} \varepsilon$. Stel nu dat $n = m+1$. Dan hebben we $w = deel\ (deel)^m(toets)^m\ toets$. Met inductie weten we dat $\mathsf{S} \overset{*}{\Rightarrow} (deel)^m(toets)^m$. Hieruit leiden we af dat het volgende geldt, wat het gewenste bewijst.

$$\mathsf{S} \overset{*}{\Rightarrow} deel\ \mathsf{S}\ toets \overset{*}{\Rightarrow} deel\ (deel)^m(toets)^m\ toets = w$$

Stel nu dat $w \in \mathcal{L}(G)$. Dan weten we dus dat $\mathsf{S} \overset{*}{\Rightarrow} w$. We bewijzen nu dat $w \in \mathcal{L}$ met inductie naar het aantal stappen n in de afleiding van $\mathsf{S} \overset{*}{\Rightarrow} w$. In het geval dat $n = 0$ dan moest gelden $w = \mathsf{S}$, maar dat is onzin. Als $n = 1$ dan geldt $w = \varepsilon$, want $w = deel\ \mathsf{S}\ \notin \Sigma^*$. Maar $\varepsilon \in \mathcal{L}$, dus dat zit goed. Nu het inductiegeval, wanneer $n = m + 1$. Dat betekent dat

$$\mathsf{S} \Rightarrow deel\ \mathsf{S}\ toets \Rightarrow deel\ v_1\ toets \Rightarrow \ldots \Rightarrow deel\ v_m\ toets = w$$

voor zekere $v_1, \ldots, v_m \in (\Sigma \cup \{\mathsf{S}\})^*$. Deze afleiding heeft m stappen, dus $v_m \in \mathcal{L}$. Daarmee geldt ook dat $w = deel\ v_m\ toets$ een element is van $\mathcal{L}$, waarmee het bewijs af is. □

Lemma 2. *De taal* $\{\ e^n o^m \mid n, m \in \mathbb{N}$ *en* $n \geq m\ \}$ *wordt herkent door de onderstaande* CFG G.

$$\mathsf{S} \to e\ \mathsf{S} \mid e\ \mathsf{S}\ o \mid \varepsilon$$

Bewijs. We bewijzen voor elke $w \in \Sigma^*$ dat $w \in \mathcal{L}(G)$ dan en slechts dan als $w \in \mathcal{L}$.

Van links-naar rechts nemen we aan dat $w \in \mathcal{L}(G)$, dus $\mathsf{S} \overset{*}{\Rightarrow} w$, en bewijzen met inductie naar het aantal afleidingsstappen n dat $w \in \mathcal{L}(G)$. Als basisgeval bekijken we $n \leq 1$, en hier is het duidelijk dat $w \in L$. De inductiestap vindt plaats voor $n = m + 1$, dus wanneer een van de onderstaande twee afleidingen gebeurt is. Hier zijn $v_1, \ldots, v_m$ rijtjes over $\Sigma \cup \{\mathsf{S}\}$. Uit de regels voor S volgt dat dit de enige twee mogelijkheden zijn.

$$\mathsf{S} \Rightarrow e\ \mathsf{S} \Rightarrow e\ v_1 \Rightarrow \ldots \Rightarrow e\ v_m = w \tag{2}$$

$$\mathsf{S} \Rightarrow e\ \mathsf{S}\ o \Rightarrow e\ v_1\ o \Rightarrow \ldots \Rightarrow e\ v_m\ o = w \tag{3}$$

In beide gevallen geldt natuurlijk dat $v_1 \Rightarrow \ldots \Rightarrow v_m$. Met inductie verkrijgen we dat $v_m \in \mathcal{L}$, dus $v_m = e^i o^j$ voor $i \geq j$. In het geval (2) zien we dat $ev_m = e^{i+1}o^j$, en omdat $i \geq j$ geldt $i + 1 \geq j$ zeker. In het geval (3) zien we $ev_m o = e^{i+1}o^{j+1}$, en wederom geldt $i + 1 \geq j + 1$. Dus in beide gevallen geldt $w \in \mathcal{L}$ zoals gewenst.

Rest ons nu de implicatie van rechts naar links aan te tonen. We nemen aan dat $w \in \mathcal{L}$, dus $w = e^i o^j$ voor $i \geq j$. Schrijf n voor het verschil $i - j$, wat niet-negatief is omdat $i \geq j$. Merk op dat $w = e^n e^j o^j$. We kunnen met inductie naar j bewijzen dat $\mathsf{S} \overset{*}{\Rightarrow} e^j o^j$. Vervolgens kun je met inductie naar n bewijzen dat $\mathsf{S} \overset{*}{\Rightarrow} e^n\mathsf{S}$. Deze twee eigenschappen samen tonen aan dat $\mathsf{S} \overset{*}{\Rightarrow} e^{j+n}o^j = e^i o^i$.

We bewijzen de eerste claim. Als $j = 0$ dan moeten we aantonen dat $\mathsf{S} \Rightarrow \varepsilon$, wat duidelijk waar is. Voor $j = j' + 1$ moeten we aantonen dat $\mathsf{S} \overset{*}{\Rightarrow} e\ e^j o^j\ o$. Met inductie weten we al dat $\mathsf{S} \overset{*}{\Rightarrow} e^j o^j$. Nu merken we op dat $\mathsf{S} \overset{*}{\Rightarrow} e\ \mathsf{S}\ o$, van waar uit het gewenste duidelijk is.

De tweede claim is te bewijzen op soortgelijke wijze. Als $n = 0$ dan moeten we aantonen dat $\mathsf{S} \overset{*}{\Rightarrow} \mathsf{S}$, wat volgt uit de reflexiviteit van $\overset{*}{\Rightarrow}$. Voor $n = n' + 1$ rest ons te bewijzen dat $\mathsf{S} \overset{*}{\Rightarrow} e\ e^n\mathsf{S}$. Dit is evenzo duidelijk, omdat we met inductie weten dat $\mathsf{S} \overset{*}{\Rightarrow} e^n\ \mathsf{S}$, dus $\mathsf{S} \overset{*}{\Rightarrow} e^{n+1}\ \mathsf{S}$ volgt al snel. □

Opgave 3 — Niet Regulier

Bewijs dat beide talen van Opgave 2 niet regulier zijn. Hint: gebruik hiervoor de Myhill–Nerode stelling of het pomplemma.

Lemma 3. *De taal* $\mathcal{L} = \{\ (deel)^n(toets)^n \mid n \in \mathbb{N}\ \}$ *is* niet *regulier.*

Bewijs met Pomplemma. Stel $\mathcal{L}$ is wel regulier. Wegens het pomplemma is er nu een pomplengte $p \in \mathbb{N}$. We kiezen nu een geniepig woord waarmee we een tegenspraak af kunnen leiden. Bekijk het mooi gekozen woord $w = (deel)^p(toets)^p$. Merk op dat $|w| = 4p + 5p = 9p \geq p$ en $w \in \mathcal{L}$.

Omdat p de pomplengte van $\mathcal{L}$ is hebben we nu $x, y, z \in \Sigma^*$ zodat $w = xyz$, $|xy| \leq p$, $|y| \geq 1$ en $xy^iz \in \mathcal{L}$ voor alle $i \in \mathbb{N}$. We merken op dat de eerste drie eigenschappen er voor zorgen dat $x = (deel)^kx'$, $y = y'(deel)^ly''$ en $z = z'(deel)^m(toets)^p$ voor zekere $k, l, m \in \mathbb{N}$ en $x', y', y'', z' \in \Sigma^*$. In het bijzonder geldt nu dat $(deel)^kx'z'$ bestaat uit *minder* dan p herhalingen van $deel$. Dit betekent dat $xz \notin \mathcal{L}$, een tegenspraak met de laatste eigenschap. Dus is $\mathcal{L}$ niet regulier. □

Bewijs met Myhill–Nerode. Bekijk het rijtje woorden $w_i = (deel)^i$ voor $i \in \mathbb{N}$. Stel $w_i \equiv w_j$. Het is duidelijk dat $w_i(toets)^i = (deel)^i(toets)^i \in \mathcal{L}$. Dus moet ook gelden dat $w_j(deel)^i = (deel)^j(toets)^i \in \mathcal{L}$. Dit dwingt af dat $i = j$. Dus de rij $w_0, w_1, \ldots$ is een oneindige rij paarsgewijs niet-equivalente woorden. Dit betekent dat $\Sigma^*/\equiv$ oneindig is, dus de Myhill–Nerode stelling vertelt ons dat $\mathcal{L}$ niet regulier is. □

Lemma 4. *De taal* $\mathcal{L} = \{\ e^no^m \mid n, m \in \mathbb{N}$ *en* $n \geq m\ \}$ *is* niet *regulier.*

Bewijs met Pomplemma. Stel $\mathcal{L}$ is regulier. Wegens het pomplemma is er een pomplengte $p \in \mathbb{N}$. Neem nu $w = e^po^p \in \mathcal{L}$ en merk op dat $|w| = 2p \geq p$. Omdat p een pomplengte is van $\mathcal{L}$ weten we dat er $x, y, z \in \Sigma^*$ zijn zo dat $xyz = w$, $|xy| \leq p$, $|y| \geq 1$ en $xy^iz \in \mathcal{L}$ voor alle $i \in \mathbb{N}$. Uit de eerste twee eigenschappen volgt dat $x = e^k$, $y = e^l$ en $z = e^mo^p$ met $k + l + m = p$. De derde eigenschap geeft ons $l \geq 1$. We zien nu dat $xz \notin \mathcal{L}$, want ware dit het geval, dan was $k + m = p$ waar, dus moet l nul zijn. Maar $0 = l \geq 1$ is onzin, dus dit geeft een tegenspraak met de vierde eigenschap. Dus $\mathcal{L}$ is niet regulier. □

Bewijs met Myhill–Nerode. Bekijk het rijtje woorden $w_i = e^i$ voor $i \in \mathbb{N}$. Stel $w_i \equiv w_j$. Neem m het maximum van i en j. Het is duidelijk dat $e^mo^m \in \mathcal{L}$, maar voor alle $l < m$ geldt $e^lo^m \notin \mathcal{L}$. Stel nu dat $i \neq j$, dan is een van de twee kleiner. Neem voor het gemak aan dat i kleiner is. Dat geeft $m = j$, dus geldt $e^io^m \in \mathcal{L}$ *zeker niet*, maar $e^jo^m \in \mathcal{L}$ is gewoon waar. Dit geeft een tegenspraak. Dus moet i gelijk zijn aan j.

Hiermee hebben we een oneidige rij $w_0, w_1, \ldots$ van paarsgewijs niet-equivalente woorden. Dit betekent dat $\Sigma^*/\equiv$ oneindig is, dus de Myhill–Nerode stelling vertelt ons dat $\mathcal{L}$ niet regulier is. □

Opgave 4 — Invullen

Bewijs het onderstaande lemma. Dit bewijs kan vrij kort, het past zeker in een pagina. Langere bewijzen worden niet geaccepteerd.

Lemma 5. *Neem twee eindige verzamelingen Σ_1 en Σ_2 en per $x \in \Sigma_1$ een context-vrije taal $\mathcal{L}_x$ over Σ_2. Stel nu dat $\mathcal{L}$ context-vrij is over Σ_1. Dan is de onderstaande taal context-vrij over Σ_2.*

$$\{\, v_1 \dots v_n \mid n \in \mathbb{N} \text{ en er is een woord } w_1 \dots w_n \in \mathcal{L} \text{ zodat } v_1 \in \mathcal{L}_{w_1}, \dots, v_n \in \mathcal{L}_{w_n} \,\}$$

Bewijs. Kies een CFG $G_x = \langle V_x, \Sigma_2, R_x, \mathsf{S}_x \rangle$ per context-vrije taal $\mathcal{L}_x$ zo dat $\mathcal{L}(G_x) = \mathcal{L}_x$. Kies ook een CFG $G = \langle V, \Sigma_1, R, \mathsf{S} \rangle$ voor $\mathcal{L}$. Nu maken we een nieuwe CFG $\underline{G} = \langle \underline{V}, \Sigma_2, \underline{R}, \mathsf{S} \rangle$. Definieer

$$\underline{V} := \bigcup_{x \in \Sigma_1} V_x \cup V,$$

dus de nieuwe variabelen worden alle oude variabelen. Definieer de nieuwe regels $\underline{R}$ als volgt. Allereerst bevallen ze alle regels uit R_x voor iedere $x \in \Sigma_1$. Verder voeg je per regel $\mathsf{A} \to w$ uit R een regel toe waarin alle letters $x \in \Sigma_1$ uit w vervangen zijn door S_x.

Stel nu dat $w \in \underline{\mathcal{L}}$. Dan betekent dat dat er een woord $v = v_1 \dots v_n \in \Sigma_1$ is zo dat $\mathsf{S} \Rightarrow v$ binnen G, en dat er woorden $u_1, \dots, u_n \in \Sigma_2^*$ zijn zo dat $\mathsf{S}_{v_i} \overset{*}{\Rightarrow} u_i$ voor alle $i = 1, \dots, n$ en $w = u_1 \dots u_n$. In onze nieuwe grammatica $\underline{G}$ is het dus zo dat $\mathsf{S} \Rightarrow \mathsf{S}_{v_1} \ \dots \ \mathsf{S}_{v_n}$, vanuit waar $\mathsf{S} \overset{*}{\Rightarrow} u_1 \dots u_n = w$ duidelijk is. Dit bewijst dat $\underline{\mathcal{L}} \subseteq \mathcal{L}(\underline{G})$

De andere kant op, als $w \in \mathcal{L}(\underline{G})$ dan geldt $\mathsf{S} \overset{*}{\Rightarrow} w$ in $\underline{G}$. Zie nu in dat als $\mathsf{S} \overset{*}{\Rightarrow} w$, dat er dan een woord $v \in \left(\bigcup_{x \in \Sigma_1} V_x\right)^*$ is zodanig dat $\mathsf{S} \overset{*}{\Rightarrow} v \overset{*}{\Rightarrow} w$. Uit onze constructie is de linker-afleiding er enkel en alleen wanneer S het met v corresponderende woord genereert. De rechter afleiding is op te breken in afleidingen $\mathsf{S}_{v_1} \Rightarrow u_1, \dots, \mathsf{S}_{v_n} \Rightarrow u_n$ waar $v = v_1 \dots v_n$ en $u_1, \dots, u_n \in \Sigma_2^*$. Dit illustreert het gewenste. □