

Automaten en Berekenbaarheid: Gekwoteerde oefenzitting 1

Prof. B. Demoen & W. Van Onsem

6 november 2014 Een oplossing

Vraag 1 (Reguliere reguliere expressies en contextvrije contextvrije grammatica's).

Is de taal van reguliere expressies over een alfabet Σ $\{r \mid r \in \text{RegEx}_\Sigma\}$ zelf regulier, of contextvrij of geen van beide? Beargumenteer je antwoord.

Is de taal van contextvrije grammatica's over een alfabet Σ en niet-terminalen NT regulier, contextvrij, of geen van beide? Beargumenteer je antwoord.

De taal van reguliere expressies over alfabet $\{a, b\}$ is contextvrij maar niet regulier. Niet-regulier omdat er geneste haakjes zijn. CF: hier zijn de regels van een CFG voor de reguliere expressies over $\{a, b\}$; het startsymbool is RE; het alfabet is $\{a, b, \phi, \epsilon, *, (,), ., | \}$

$\text{RE} \rightarrow \epsilon$

$\text{RE} \rightarrow \phi$

$\text{RE} \rightarrow a$

$\text{RE} \rightarrow b$

$\text{RE} \rightarrow (\text{RE})^*$

$\text{RE} \rightarrow \text{RE} . \text{RE}$

$\text{RE} \rightarrow \text{RE} \mid \text{RE}$

Dit kan gemakkelijk naar een willekeurig alfabet veralgemeend worden.

De verzameling regels van een contextvrije grammatica's alfabet $\{a, b\}$ en niet-terminalen $\{S, A, B\}$ is regulier. Hier is een reguliere expressie voor de regels:

$((S \mid A \mid B) \rightarrow (a \mid b \mid S \mid A \mid B)^*)^*$

Je kan hiervan vertrekken om voor het hele CFG 4-tal een reguliere expressie op te stellen, maar de regels zijn de essentie.

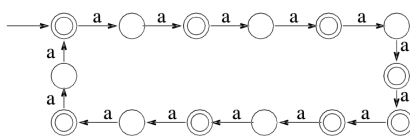
Pas toch op: de $\rightarrow$ en $*$ in de CFG voor RE, hebben echt een andere betekenis dan de $\rightarrow$ en $*$ in de RegExp voor de regels van een CFG ...

Vraag 2. Geef twee bewijzen van de volgende uitspraak: er bestaat geen DFA met 5 toestanden die de taal $\{a^n \mid n \text{ is deelbaar door } 2 \text{ of } 7\}$ beslist. (de *of* hiervoor is niet exclusief)

Je eerste bewijs steunt op de theorie van minimale DFAs, je tweede bewijs steunt op het pompend lemma. Het is natuurlijk mogelijk dat je nog andere stukjes uit de cursus kan gebruiken.

Je verdient een **bonus** met een goed bewijs dat op Myhill-Nerode relaties steunt, maar enkel als je de eerste twee bewijzen goed hebt.

Bewijs met minimale DFA:



Figuur 1: ...

Deze DFA is minimaal omdat elke twee toestanden f-verschillend zijn (dat moet je natuurlijk kunnen argumenteren) en alle toestanden zijn bereikbaar. Hij heeft 14 toestanden, dus een DFA voor dezelfde taal met 5 toestanden kan niet.

Bewijs met het pompend lemma: Het bewijs van het PL zelf geeft als pomplengte één meer dan het aantal toestanden van de DFA. Stel dus dat er een DFA bestaat met 5 toestanden voor die taal. Dan is 6 een goede pomplengte: PL garandeert nu dat elke string niet korter dan 6 en die tot de taal behoort, een opdeling heeft die gepompt kan worden (en aan de drie voorwaarden van het PL voldoet). Neem a^7 : die behoort tot de taal. Nu zie je dat elke gepaste opdeling door pompen een string kan opleveren die niet tot de taal behoort (daarbij hoort iets meer uitleg om perfect te zijn).

Bewijs met MN(L): niemand verdiende de bonus - de vraag kan in de toekomst gerecycleerd worden.

Vraag 3 (Pompend lemma). Bewijs met behulp van het pompend lemma dat $L = \{0^n 1^m 2^{n+m} | n, m \geq 0\}$ niet regulier is. Indien L contextvrij is, geef dan een contextvrije grammatica, of probeer met het pompend lemma voor contextvrije talen te bewijzen dat L niet contextvrij is.

L is niet regulier: stel er bestaat een pomplengte p voor L , beschouw dan de string $0^p 1^{2p} 2^{3p}$. Alle manieren van opdelen xyz zodat xy^iz ook in de taal moeten zitten, worden gedekt door de volgende 2 gevallen:

y bestaat enkel uit één soort karakters: oppompen levert dan een string met te veel van dat karakter

y bestaat minstens twee verschillende karakters: oppompen levert dan een string met een verkeerde volgorde van karakters

Een alternatief bewijs (zonder het PL): neem de doorsnede van L met de reguliere taal bepaald door de RegExp $(01)^*$. Die doorsnede is gelijk aan $\{0^n 1^n | n \geq 0\}$ en is niet regulier. Vermits de doorsnede van reguliere talen regulier is, kan L niet regulier zijn.

Een CFG voor L is $(\{0, 1, 2\}, R, S)$ waarbij R gelijk is aan

$S \rightarrow 0S2$

$S \rightarrow A$

$A \rightarrow 1A2$

$A \rightarrow \epsilon$

Vraag 4 (Chomsky normaal vorm). Geef voor de taal $\{a^n b^n | n \geq 0\}$ een CFG in Chomsky normaal vorm. Laat zien hoe je dat deed!

Vertrek van een voor de hand liggende CFG voor die taal - de regels zijn

$S \rightarrow \epsilon$ (1)

$S \rightarrow aSb$ (2)

Die transformeren we volgens het algoritme in de cursus: (1) mag blijven; (2) wordt vervangen door

$S \rightarrow AB$ (3)

$S \rightarrow AXB$ (4)

$X \rightarrow AXB$ (5)

$A \rightarrow a$ (6)

$B \rightarrow b$ (7)

(3), (6) en (7) zijn al ok, en (4) en (5) vervangen we door

$S \rightarrow YB$

$X \rightarrow YB$

$Y \rightarrow AX$

en we zijn klaar.

Vraag 5 (Algebra van talen). Hieronder wordt gevraagd naar voorbeelden van talen/strings die samen aan een aantal voorwaarden voldoen: verantwoord je voorbeeld. Indien je denkt dat zulk een voorbeeld niet bestaat, dan argumenteer je waarom. We gebruiken de uitdrukking L is *echt CF* als afkorting voor $L \in CFL \setminus RegLan$.

1. L_1 en L_2 , beide echt CF en $L_1 \cup L_2 \in RegLan$

dit kan: neem een CFL waarvan het complement ook CF is (die is dan ook echt CF, want kan niet regulier zijn als complement van een echte CF); de unie is de reguliere taal Σ^* ; met $L_1 = \{a^n b^n | n \geq 0\}$ moet dat lukken - je moet wel de nodige eigenschappen aantonen; een ander concreet voorbeeld: $L_1 = \{a^n b^m | n = m\}$ en $L_2 = \{a^n b^m | n \neq m\}$; de unie is een taal met reguliere expressie $a^* b^*$.

2. L_1 en L_2 , beide echt CF en $L_1 \cap L_2 \in RegLan$

dit kan: neem L_1 en L_2 zoals hierboven; de doorsnede is leeg, dus regulier; een ander voorbeeld: $L_1 = \{a^n b^n | n \geq 0\}$ en $L_2 = \{b^n a^n | n \geq 0\}$.

3. L echt CF, $s \in \Sigma^*$ en $L \cap \{s\} \in RegLan$

dit is altijd waar: als s in L zit dan is $L \cap \{s\} = \{s\}$; als s niet in L zit, dan is $L \cap \{s\} = \emptyset$; in beide gevallen is dat een eindige verzameling en dus regulier

Vraag 6 (Myhill-Nerode relaties). Gegeven de verzameling $V = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ en twee partities van V : $A = \{\{0, 1\}, \{2, 3\}, \{4, 5\}, \{6, 7\}, \{8, 9\}\}$ en $B = \{\{0, 2\}, \{1, 4\}, \{3\}, \{5, 9\}, \{6, 7\}, \{8\}\}$. Geef niet enkel het resultaat maar ook de tussenstappen die je deed, of de redenering die je volgde.
Bepaal

(a) het supremum van A en B

0 moet in een element X van het supremum zitten - laat ons X opbouwen:

vermits $0 \in X$, moeten ook 1 (wegens de $\{0, 1\}$ in A) en 2 (wegens de $\{0, 2\}$ in B) in X zitten

vermits $1 \in X$ moet ook 4 $\in X$ (zie B)

vermits $2 \in X$ moet ook 3 $\in X$ (zie A)

$3 \in X$ triggered geen nieuwe elementen voor X

4 impliceert 5 ($\{4, 5\}$ in A)

5 impliceert 9 ($\{5, 9\}$ in B)

9 impliceert 8

8 impliceert niks extra meer

dus $X = \{0, 1, 2, 3, 4, 5, 8, 9\}$. Ook $\{6, 7\}$ is een element van het supremum.

Alles in nu afgedekt, dus het supremum is $\{\{0, 1, 2, 3, 4, 5, 8, 9\}, \{6, 7\}\}$.

(b) het infimum van A en B

In het infimum zitten maximale elementen die zelf behoren tot een element van A en een element van B . We kunnen opnieuw beginnen met een element X van het infimum te bepalen waarin 0 zit.

1 kan niet in X zitten, want $\{0, 1\}$ is geen deel van een element van B

2 kan niet in X zitten, want $\{0, 2\}$ is geen deel van een element van B

...

dus X is gelijk aan $\{0\}$

Diezelfde redenering kan je volgen voor 1, 2, 3, 4, 5, 8, 9.

Tenslotte, als Y in het infimum zit en 6 bevat, dan moet 7 ook in Y , want anders was Y niet maximaal.

Het infimum is dus $\{0\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{8\}, \{9\}, \{6, 7\}$