TENTAMEN

Basismodellen in de Informatica

vakcode: 211180
datum: 2 juli 2009
tijd: 9:00–12:30 uur

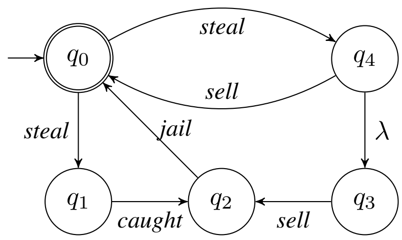
<u>VOORBEELDUITWERKING</u>

Algemeen

- Bij dit tentamen mag gebruik worden gemaakt van het boek van Sudkamp, van de handleiding van Basismodellen in de Informatica, en van de handouts van de colleges.
- Dit tentamen bestaat uit 6 opgaven, waarvoor in het totaal 95 punten behaald kunnen worden. Het minimale aantal punten per opgave bedraagt 0. Het cijfer is het aantal punten gedeeld door 10, afgerond tot een geheel getal, plus 1. (Er kunnen dus 5 bonuspunten worden verdiend.)

Opgave 1 **(20 punten)**

Beschouw de volgende NFA-λ, M, die de pakkans van Jim de Burglar in kaart brengt:



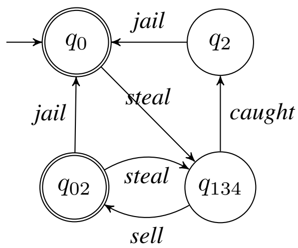
a. Beschrijf de "input transitie-functie" van M in tabelvorm.
b. Transformeer M naar een equivalente DFA, gebruikmakend van de subset-constructie.

Antwoord op Opgave 1

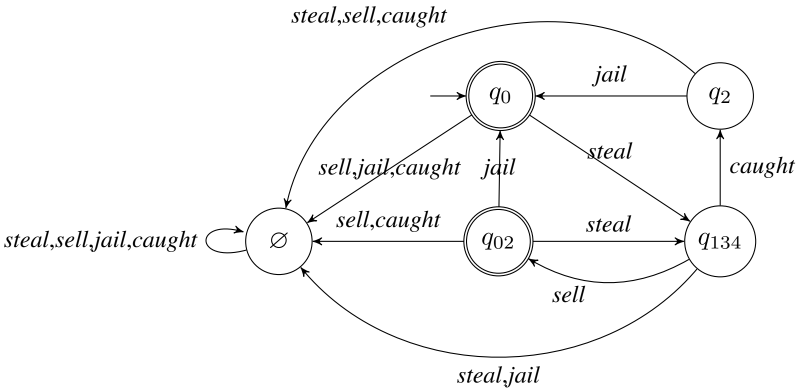
a. *(8 punten)*

	$\lambda-$closure	*steal*	*sell*	*caught*	*jail*
q_0	$\{q_0\}$	$\{q_1, q_3, q_4\}$	$\varnothing$	$\varnothing$	$\varnothing$
q_1	$\{q_1\}$	$\varnothing$	$\varnothing$	$\{q_2\}$	$\varnothing$
q_2	$\{q_2\}$	$\varnothing$	$\varnothing$	$\varnothing$	$\{q_0\}$
q_3	$\{q_3\}$	$\varnothing$	$\{q_2\}$	$\varnothing$	$\varnothing$
q_4	$\{q_3, q_4\}$	$\varnothing$	$\{q_0, q_2\}$	$\varnothing$	$\varnothing$

b. *(12 punten)*



or complete:



Opgave 2 (20 punten)

Beschouw de volgende twee talen:

$$\begin{array}{lcl}\mathcal{L}_1 & := & \{a^{2m}b^{3m} \mid m > 0\} \\ \mathcal{L}_2 & := & \{w \in \{a,b\}^* \mid 3 \cdot \#_a(w) = 2 \cdot \#_b(w)\}\end{array}$$

waarbij $\#_a(w)$ het aantal voorkomens van a in w betekent.

a. Bewijs met behulp van de pompstelling dat $\mathcal{L}_1$ niet regulier is.

b. Laat zien dat dan ook $\mathcal{L}_2$ niet regulier is. Gebruik de geslotenheidseigenschappen voor reguliere talen.

c. Is $\mathcal{L}_1$ contextvrij? Toon uw antwoord aan.

d. Geef een Turingmachine die $\mathcal{L}_2$ beslist.

Antwoord op Opgave 2

a. *(5 punten)* $\mathcal{L}_1 = \{a^{2m}b^{3m} \mid 0 \leq i \leq j\}$. Stel $\mathcal{L}_1$ is regulier. Laat $k > 0$ gegeven, kies $z = a^{2k}b^{3k}$, dan $|z| = 5k > k$, en $z \in \mathcal{L}_1$. Stel $z = uvw$ met $|uv| \leq k$ en $|v| \geq 1$. Laat $z' = uv^2w$. Volgens de pompstelling geldt dan $z' \in \mathcal{L}_1$.

   Echter, $u = a^p$, $v = a^q$ en $w = a^{2k-p-q}b^{3k}$ voor zekere p, q, met $q > 0$. dus $z' = uv^2w = a^{2k+q}b^{3k}$. Omdat $q > 0$, geldt $z' \notin \mathcal{L}_1$. Tegenspraak. Dus $\mathcal{L}_1$ is niet regulier.

b. *(5 punten)* Stel $\mathcal{L}_2$ is regulier. Merk op dat $\mathcal{L}_1 = \mathcal{L}_2 \cap a^+b^+$. Reguliere talen zijn gesloten onder doorsnede, dus $\mathcal{L}_1$ is dan regulier. Tegenspraak. Dus $\mathcal{L}_2$ is niet regulier.

   N.B.: hierboven a^*b^* gebruiken is niet (helemaal) goed.

c. *(5 punten)* Ja, want er is een contextvrije grammatica: $S \to aabbb \mid aaSbbb$

d. *(5 punten)*

Opgave 3 (15 punten)

Definieer: v is een deelwoord van w als je v verkrijgt door letters uit w te schrappen. De taal van alle deelwoorden van L kan als volgt worden gedefinieerd:

$$DW(L) = \{a_1 \cdots a_n \mid w_0 a_1 w_1 \cdots w_{n-1} a_n w_n \in L\}$$

Bewijs nu de volgende stelling: Als L een reguliere taal is, dan is ook $DW(L)$ regulier.

Antwoord op Opgave 3

N.B.: in het oorspronkelijke tentamen zat er een spelfout in de definitie; deze was namelijk:

$$DW(L) = \{a_1 \cdot a_n \mid w_0 a_1 w_1 \cdots w_{n-1} a_n w_n \in L\}$$

Als L regulier is dan is er een NFA-λ, zeg M, zodat $\mathcal{L}(M) = L$. Laat M' een NFA-λ zijn, die voor elke pijl (s, a, t) in M een pijl (s, λ, t) toevoegt. De berekeningen van M' zijn dan precies de berekenigen van M, waarbij sommige letters door λ zijn vervangen. Dus $\mathcal{L}(M') = DW(L)$.

Opgave 4 (15 punten)

Gegeven is de volgende contextvrije grammatica G:

$$\begin{aligned} S &\rightarrow ASB \mid BA \\ A &\rightarrow aaA \mid a \\ B &\rightarrow bbB \mid \lambda \end{aligned}$$

a. Transformeer G stapsgewijs naar Chomsky normaalvorm.
b. Gebruik het CYK algoritme om het woord $w = aabb$ te ontleden.
c. Laat zien hoe de ontleedboom van w uit de matrix is af te leiden.

Antwoord op Opgave 4

a. *(7 punten)* Verwijderen startrecursie: voer een nieuw startsymbool S' in

$$\begin{aligned} S' &\rightarrow S \\ S &\rightarrow ASB \mid BA \\ A &\rightarrow aaA \mid a \\ B &\rightarrow bbB \mid \lambda \end{aligned}$$

Na verwijderen λ-regels:

$$\begin{aligned} S' &\rightarrow ASB \mid BA \mid AS \mid A \\ S &\rightarrow ASB \mid BA \mid AS \mid A \\ A &\rightarrow aaA \mid a \\ B &\rightarrow bbB \mid bb \end{aligned}$$

Na verwijderen ketting-regels:

$$\begin{aligned} S' &\rightarrow ASB \mid BA \mid AS \mid aaA \mid a \\ S &\rightarrow ASB \mid BA \mid AS \mid aaA \mid a \\ A &\rightarrow aaA \mid a \\ B &\rightarrow bbB \mid bb \end{aligned}$$

Fatsoeneren volgens het vaste stramien (links) of door de twee a's en de twee b's te combineren (rechts):

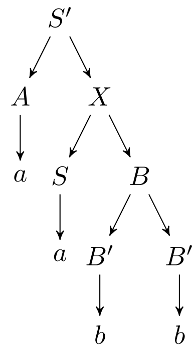
$$
\begin{aligned}
S' &\to AX \mid BA \mid AS \mid A'Y \mid a \\
S &\to AX \mid BA \mid AS \mid A'Y \mid a \\
A &\to A'Y \mid a \\
B &\to B'Z \mid B'B' \\
X &\to SB \\
Y &\to A'A \\
Z &\to B'B \\
A' &\to a \\
B' &\to b
\end{aligned}
\qquad\qquad
\begin{aligned}
S' &\to AX \mid BA \mid AS \mid A''A \mid a \\
S &\to AX \mid BA \mid AS \mid A''A \mid a \\
A &\to A''A \mid a \\
B &\to B''B \mid B'B' \\
X &\to SB \\
A'' &\to A'A' \\
B'' &\to B'B' \\
A' &\to a \\
B' &\to b
\end{aligned}
$$

b. *(5 punten)* Volgens de twee grammatica's hierboven:

	1	2	3	4
1	S, S', A, A'	S, S', Y	–	S, S', X
2		S, S', A, A'	–	X
3			B'	B
4				B'

	1	2	3	4
1	S, S', A, A'	S, S', A''	–	S, S', X
2		S, S', A, A'	–	X
3			B'	B, B''
4				B'

c. *(3 punten)* De afleidingsbomen is voor beide grammatica's identiek:



Opgave 5 (15 punten)

Beschouw de volgende onbeperkte grammatica G:

$$
\begin{aligned}
S &\to ASB \mid \lambda \\
AB &\to ba \mid Ba \\
Ab &\to bA \\
Aa &\to aA \mid a
\end{aligned}
$$

a. Geef een afleiding van $baba$

b. Leg uit waarom $\#_a(w) = \#_b(w)$ voor alle $w \in L(G)$

c. Welke taal genereert deze grammatica? Beredeneer je antwoord zo goed mogelijk.

d. Geef een equivalente contextvrije grammatica in Greibach normaalvorm, met zo weinig mogelijk veriabelen.

Antwoord op Opgave 5

a. *(2 punten)* $S \Rightarrow ASB \Rightarrow AASBB \Rightarrow AABB \Rightarrow AbaB \Rightarrow bAaB \Rightarrow baAB \Rightarrow baba$

b. *(5 punten)* Dit is als volgt in te zien:

- De regel voor S genereert de taal $\{A^iB^i \mid i \geq 0\}$
- Om een B kwijt te raken moet de regel $AB \rightarrow ba$ worden toegepast. Woorden met minder A's dan B's kunnen daarom nooit termineren. (De regels $AB \rightarrow Ba$ en $Aa \rightarrow a$ kunnen dus niet worden toegepast in een terminerende afleiding.)
- Hieruit volgt dat $\#_A(w) + \#_a(w) = \#_B(w) + \#_b(w)$ voor alle zinsvormen $w \in \{A, B, a, b\}^*$ met $w \in L(G)$. (Formeel kan dit worden bewezen met inductie over de lengte van de afleiding $A^iB^i \Rightarrow w$.)

c. *(4 punten)* De taal is $(ba)^*$ (een reguliere taal). Dit komt doordat a's en b's "bevroren" worden in de positie waarin ze ontstaan.

d. *(4 punten)* De voor de hand liggende reguliere grammatica is in Greibach normaalvorm. Een alternatief is rechts gegeven. Kleiner kan niet, aangezien S niet rechts mag voorkomen, en regels zoals $T \rightarrow baT$ niet toegestaan zijn.

$$\begin{array}{ll} S \rightarrow bA \mid \lambda & \qquad S \rightarrow bAT \mid bA \mid \lambda \\ A \rightarrow aB \mid a & \qquad T \rightarrow bAT \mid bA \\ B \rightarrow bA & \qquad A \rightarrow a \end{array}$$

Opgave 6 **(15 punten)**

Zij R een injectieve functie die een willekeurige Turingmachine M omzet in een string-codering hiervan; d.w.z., $R(M) \in \Sigma^*$ voor een gegeven alfabet Σ, en $M_1 \neq M_2$ impliceert $R(M_1) \neq R(M_2)$. Zij

$$L_1 = \{R(M) \mid M \text{ is een willekeurige Turingmachine}\}$$
$$L_2 = \{R(M)w \mid M \text{ is een Turingmachine die stopt voor invoer } w\}$$

Beantwoord de volgende vragen, door te beredeneren hoe zo'n grammatica er uit zou zien of waarom zo'n grammatica niet kan bestaan. Kies eventueel zelf een geschikte codering R (die aan bovenstaande voorwaarden voldoet) zodat het antwoord eenvoudiger wordt.

a. Bestaat er een contextgevoelige grammatica die de taal L_1 genereert?
b. Bestaat er een contextgevoelige grammatica die de taal L_2 genereert?
c. Bestaat er een onbeperkte grammatica die de taal L_1 genereert?
d. Bestaat er een onbeperkte grammatica die de taal L_2 genereert?

Antwoord op Opgave 6

a. *Bestaat er een contextgevoelige grammatica die de taal L_1 genereert? (4 punten)*

Ja, zo'n grammatica bestaat. R kan bij voorbeeld eerst de toestanden op een rij zetten, dan de overgangen tussen deze toestanden, en tenslotte de eindtoestanden. Het is niet moeilijk een lineair begrensde Turingmachine te bouwen die nagaat of een gegeven string aan dit formaat voldoet. Dit impliceert dat er ook een contextgevoelige grammatica bestaat die de taal genereert.

b. *Bestaat er een contextgevoelige grammatica die de taal L_2 genereert? (4 punten)*

Nee, zo'n grammatica bestaat niet. Als hij zou bestaan, zou de taal recursief zijn (want dan bestaat er ook een linear begrensde Turingmachine die de taal beslist); dat druist in tegen de onbeslisbaarheid van het halting-probleem. (Zie Corollary 12.1.2 in het boek.)

c. *Bestaat er een onbeperkte grammatica die de taal L_1 genereert? (3 punten)*

Ja, zo'n grammatica bestaat, want er bestaat al een contextgevoelige grammatica (en dat is volgens de Chomsky-hiërarchie een kleinere klasse dan de onbeperkte grammatica's).

d. *Bestaat er een onbeperkte grammatica die de taal L_2 genereert? (4 punten)*

Ja, zo'n grammatica bestaat, want er bestaat een universele Turingmachine die deze taal herkent, en onbeperkte grammatica's zijn (volgens de Chomsky-hiërarchi) even krachtig als Turingmachines.