Tentamen Automatentheorie en Formele Talen

(met uitwerkingen en toelichting)

Vakcode 2M130, 29 juni 2001, 9.00 - 12.00 uur

Dit tentamen bestaat uit vijf opgaven die alle even zwaar tellen. Daarnaast kan met de inleveropgaven een bonus van maximaal een punt behaald worden, met dien verstande dat het eindcijfer niet boven de tien uitkomt. Wie voor een bonus in aanmerking wil komen wordt verzocht de naam van de instructieleider te vermelden.

Voor alle vragen geldt: motiveer uw antwoord.

Opgave 1.

Gegeven is de taal

$$L = L(a^*b(a+b)b^*)$$

over het alfabet $\{a, b\}$.

a. Geef een nfa M met $L(M) = L$.

b. Geef een nfa N met $L(N) = \overline{L}$.

Toelichting en uitwerking:
Onderdeel a 4 punten.
Onderdeel b 6 punten.

Het is niet nodig precies het algoritme uit het boek om een reguliere expressie te vertalen naar een nfa precies te volgen; dit genereert allerlei overbodige λ-stappen. Zo is een goed antwoord op a een nfa met drie toestanden q_0, q_1, q_2 met $F = \{q_2\}$ en a-stappen van q_0 naar q_0 en van q_1 naar q_2, en b-stappen van q_0 naar q_1 van q_1 naar q_2, en van q_2 naar q_2. Voor het antwoord op b kun je hier eerst een dfa van maken door een trap state q_3 toe te voegen en een a-stap van q_2 naar q_3 en zowel een a-stap als een b-stap van q_3 naar zichzelf. Als je in deze dfa kiest $F = \{q_0, q_1, q_3\}$ heb je een goed antwoord op b.

Zowel een nfa als een dfa kan gegeven worden door alleen de transitiegraaf te tekenen; de opsomming van alle ingrediënten van het zoveel-tupel kan dan impliciet worden gelaten.

Opgave 2.

Gegeven is de taal

$$L = \{a^n b^m a^n \mid m, n \geq 0\}.$$

a. Geef een contextvrije grammatica G met $L(G) = L$.

b. Bewijs dat L niet regulier is.

Toelichting en uitwerking:
Onderdeel a 4 punten. Antwoord:

$$S \to aSa|A,\ A \to bA|\lambda$$

Onderdeel b 6 punten. Voor de hand liggende oplossing is met de pompstelling:

Stel dat L regulier is. Dan is er volgens de pompstelling een getal m zodanig dat er voor elke w met $|w| \geq m$ een schrijfwijze $w = xyz$ bestaat waarin $|xy| \leq m$ en $|y| \geq 1$ en waarvoor $xy^iz \in L$ voor elke $i \geq 0$. Kies nu $w = a^m b^m a^m$. Vanwege $|xy| \leq m$ en $|y| \geq 1$ is nu $y = a^k$ voor $k = |y| \geq 1$ en $xy^0z = xz = a^{m-k}b^m a^m \notin L$. Dit is in tegenspraak met de pompstelling, dus is L niet regulier.

Als de precieze keuze voor w ontbreekt kun je voor dit onderdeel maximaal 1 punt krijgen (tenzij er een ander correct bewijs gegeven wordt). Als de keuze voor w wel gegeven wordt maar de rest van de redenering niet goed is kun je maximaal 3 punten voor dit onderdeel krijgen.

Opgave 3.

Geef een non-deterministische pushdown-automaat die precies de taal beschreven door de volgende grammatica accepteert:

$$\begin{array}{lcl} S & \to & aSbb \mid bA \\ A & \to & a \mid bSa. \end{array}$$

Toelichting en uitwerking:
Deze grammatica is makkelijk naar Greibach normaalvorm

$$\begin{array}{lcl} S & \to & aSBB \mid bA \\ A & \to & a \mid bSC \\ B & \to & b \\ C & \to & a \end{array}$$

te brengen. Let wel:

$$\begin{array}{lcl} S & \to & aSBB \mid bA \\ A & \to & a \mid bSA \\ B & \to & b \end{array}$$

is niet goed: deze genereert ook *bbbabbaa* terwijl die niet door de gegeven grammatica wordt gegenereerd.

Vanuit de Greibach normaalvorm is er een rechtstreekse vertaling naar een pushdown-automaat:

$$M = (\{q_0, q_1, q_2\}, \{a, b\}, \{A, B, C, S, z\}, \delta, q_0, z, \{q_2\})$$

met

$$\delta(q_0, \lambda, z) = \{(q_1, Sz)\},\quad \delta(q_1, \lambda, z) = \{(q_2, z)\},$$

$$\delta(q_1, a, S) = \{(q_1, SBB)\},\quad \delta(q_1, b, S) = \{(q_1, A)\},$$

$$\delta(q_1, a, A) = \{(q_1, \lambda)\}, \quad \delta(q_1, b, A) = \{(q_1, SC)\},$$
$$\delta(q_1, a, C) = \{(q_1, \lambda)\}, \quad \delta(q_1, b, B) = \{(q_1, \lambda)\},$$

en $\delta(\cdots) = \emptyset$ voor alle andere gevallen.

Als bovenstaande foutieve Greibach normaalvorm wordt gebruikt maar daarvoor wel de goede vertaling naar een pushdown-automaat is gegeven kun je maximaal 7 punten krijgen.

Het is ook mogelijk een rechtstreekse vertaling te geven zonder Greibach normaalvorm als tussenformaat. Dan moet je zowel variabelen als terminals op de stack zetten.

Opgave 4.

Welke van de volgende beweringen zijn waar? Geef van de ware beweringen een bewijs (eventueel gebruikmakend van stellingen uit het boek) en van de niet-ware een tegenvoorbeeld.

a. Als L_1 en L_2 contextvrij zijn dan is $L_1L_2 \cup L_2L_1$ ook contextvrij.

b. Als L_1 en L_2 regulier zijn dan is $L_1L_2 \cup L_2L_1$ ook regulier.

c. Als L_1 en L_2 contextvrij zijn dan is $L_1^R \cap L_2$ ook contextvrij.

d. Als L_1 en L_2 regulier zijn dan is $L_1^R \cap L_2$ ook regulier.

Toelichting en uitwerking:

Beweringen a is waar volgens de stelling die zegt dat de klasse van contextvrije talen gesloten is onder vereniging en concatenatie. Beweringen b en d zijn waar volgens de stelling die zegt dat de klasse van reguliere talen gesloten is onder doorsnede, vereniging, concatenatie en reverse.

Bewering c is niet waar: kies bijv. $L_1 = \{c^k b^n a^n | k, n \geq 0\}$ en $L_2 = \{a^k b^n c^n | k, n \geq 0\}$. Dan is $L_1^R \cap L_2 = \{a^n b^n c^n | n \geq 0\}$. Zonder bewijs mag bekend worden verondersteld dat $L_1^R \cap L_2$ niet contextvrij is; dit is een standaardvoorbeeld. De taal L_1 wordt gegenereerd door de grammatica

$$S \to cS|A, \quad A \to \lambda|bAa,$$

en is dus contextvrij; de taal L_2 wordt gegenereerd door de grammatica

$$S \to aS|A, \quad A \to \lambda|bAc,$$

en is dus ook contextvrij.

Puntentelling: a, b en d elk twee punten als een goede motivering gegeven wordt, bijvoorbeeld verwijzing naar genoemde geslotenheidsstellingen, elk een punt als het antwoord 'waar' wordt gegeven zonder goede motivering. Onderdeel c levert 1 punt op als het antwoord 'onwaar' wordt gegeven zonder goede motivering; 3 punten als goede tegenvoorbeelden worden gegeven zonder goede motivering en de volle 4 punten als voor L_1 en L_2 een bijbehorende grammatica of pushdown-automaat wordt gegeven.

Opgave 5.

Gegeven is de grammatica

$$\begin{aligned} S &\rightarrow aX \mid abYaba \\ X &\rightarrow \lambda \mid bYa \\ Y &\rightarrow \lambda \mid aXb \end{aligned}$$

a. Toon aan dat deze grammatica ambigu is.

b. Geef een equivalente grammatica met slechts één variabele die niet ambigu is.

Toelichting en uitwerking:

Onderdeel a 4 punten. Je kunt laten zien dat *ababa* twee verschillende ontleedbomen heeft:

```
  S
 / \
a   X
  / | \                S
 b  Y  a
  / | \            / / / \ \ \
 a  X  b          a b Y   a b a
    |                 |
    λ                 λ
```

Als er niet expliciet zo'n string wordt gegeven maar alleen naar de definitie van ambiguiteit wordt gerefereerd worden geen punten voor dit onderdeel geven. Als iemand wel zo'n string geeft maar daarvan alleen opmerkt dat die verschillende afleidingen heeft worden maximaal 2 punten gegeven. Het geven van de twee ontleedbomen is genoeg om de 4 punten te krijgen, veel meer is er ook niet over te zeggen.

Onderdeel b 6 punten. Je kunt opmerken (zonder gedetailleerd bewijs) dat de taal $L(a(ba)^*)$ is en daarvoor de grammatica

$$S \rightarrow a|Sba$$

geven, of het spiegelbeeld.

Een meer stapsgewijze aanpak vanuit de gegeven grammatica is:

1. verwijder eerst de overbodige productie $S \rightarrow abYaba$;

2. vervang $X \rightarrow bYa$ door $X \rightarrow ba|baXba$ volgens de standaardsubstitutieregel;

3. verwijder nu de overbodige variabele Y en de bijbehorende producties; we hebben nu:

$$\begin{aligned} S &\rightarrow aX \\ X &\rightarrow \lambda \mid ba \mid baXba \end{aligned}$$

4. vervang dit door

$$\begin{aligned} S &\rightarrow aX \\ X &\rightarrow \lambda \mid ba \mid bSba \end{aligned}$$

5. de standaardsubstitutieregel levert nu

$$S \rightarrow a \mid aba \mid abSba$$

Deze heeft inderdaad slechts één variabele en is niet ambigu.