TENTAMEN AUTOMATA THEORY
Maandag 19 december 2022, 09.00 - 12.00 uur

Dit tentamen bestaat uit acht opgaven, waarbij steeds tussen [en] staat hoeveel punten er ongeveer mee te verdienen zijn. In totaal zijn er 100 punten te verdienen. Wanneer er bij een vraag om uitleg, motivatie of toelichting gevraagd wordt, is het belangrijk om die ook te geven.
Als we het in dit tentamen over een eindige automaat hebben (zonder verdere toevoeging), bedoelen we een deterministische eindige automaat zonder Λ-transities (wat elders *DFA* genoemd wordt).

---

1. [8 pt] Laat

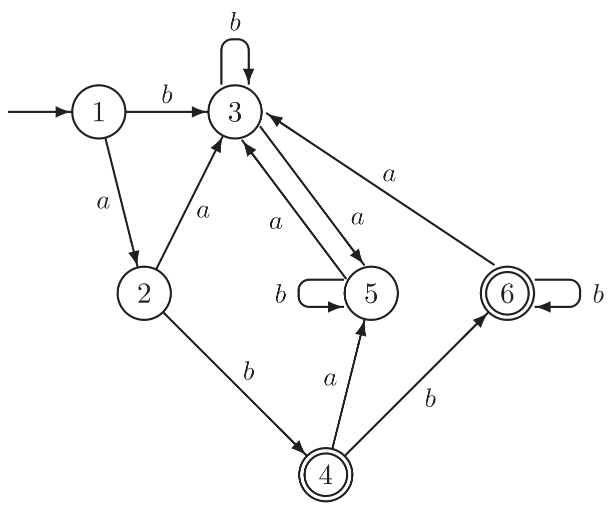
$$L = \{x \in \{a,b\}^* \quad | \quad x \text{ begint met } aa \text{ of } ab\text{, en } x \text{ eindigt met } aa \text{ of } bb\}$$

Dus bijvoorbeeld $aa \in L$, $abb \in L$, maar $abbbab \notin L$.

Teken een eindige automaat M, zó dat $L(M) = L$.

---

2. [9 pt] Beschouw de volgende eindige automaat M_1:



Pas op M_1 het algoritme toe om een minimale eindige automaat M_2 (een eindige automaat met zo weinig mogelijk toestanden) te vinden, zó dat $L(M_2) = L(M_1)$. Dat wil zeggen:

(a) Loop herhaaldelijk, kolom voor kolom (kolommen van links naar rechts), door onderstaande driehoek, en vul die in met cijfers die aangeven in welke iteratie van het algoritme is vastgesteld dat toestanden i en j niet kunnen worden samengevoegd.

$i = 2$	·				
3	·	·			
4	·	·	·		
5	·	·	·	·	
6	·	·	·	·	·
	$j =$ 1	2	3	4	5

Geef als je antwoord de resulterende driehoek.

(b) Teken de eindige automaat M_2 die uit M_1 ontstaat door toestanden die volgens je antwoord bij onderdeel (a) kunnen worden samengevoegd, daadwerkelijk samen te voegen. Uiteraard moeten de namen van de toestanden van M_1 herkenbaar zijn in de namen van de toestanden van M_2.

---

3. [17 pt]

(a) Laat $M_1 = (Q_1, \Sigma, q_1, A_1, \delta_1)$ een willekeurige eindige automaat zijn, en laat $L_1 = L(M_1)$.

Hoe kun je M_1 ombouwen tot een eindige automaat $M_2 = (Q_2, \Sigma, q_2, A_2, \delta_2)$, zó dat $L(M_2) = L_1'$ (het complement van L_1)?

Geef als je antwoord de vier onderdelen Q_2, q_2, A_2 en δ_2 van M_2, uitgedrukt in de onderdelen van M_1.

Als je het antwoord op dit onderdeel niet weet, dan kun je het 'kopen' van de docent. Wellicht kun je dan wel onderdeel (b) maken.

(b) De constructie uit onderdeel (a) is niet over te zetten naar willekeurige **stapel**automaten. Geef twee redenen waarom je niet op dezelfde manier een willekeurige stapelautomaat M_1 om kunt bouwen tot een stapelautomaat M_2, zó dat $L(M_2) = L(M_1)'$.

Illustreer elk van de redenen met een concrete voorbeeld stapelautomaat M_1 met hoogstens vier toestanden, waarin de constructie een automaat M_2 oplevert met $L(M_2) \neq L(M_1)'$. Geef bij elke voorbeeld stapelautomaat ook een concrete string x, zó dat x door zowel M_1 als M_2 geaccepteerd wordt, of juist door beide automaten niet geaccepteerd wordt.

(c) Inderdaad is de klasse van context-vrije talen niet gesloten onder complement. Geef een voorbeeld van een context-vrije taal L_1 waarvan het complement L_1' niet context-vrij is.

Als je een voorbeeldtaal L_1 noemt, waarvan in de hoorcolleges is uitgelegd dat L_1 context-vrij is en L_1' niet, hoef je dat verder niet aan te tonen. Anders moet je dat wel aantonen.

---

4. [12 pt] Deze opgave gaat over reguliere talen over het alfabet $\{a, b, c\}$.
Laat $L_1 = \{a^i b^j c^k \mid i, j, k \geq 1\}$.

(a) Geef een reguliere expressie voor de taal L_1.

(b) Geef een reguliere expressie voor de taal L_1' (het complement van L_1).

Als je geen geschikte reguliere expressie kunt bedenken, kun je een deel van de punten voor dit onderdeel verdienen met een reguliere expressie voor de taal L_0' (het complement van L_0), waarbij $L_0 = \{a^i b^j c^k \mid i, j, k \geq 0\}$. Vermeld in dat geval expliciet dat je voor L_0' gekozen hebt.

Leg ook uit waarom je expressie precies de gekozen taal beschrijft.

---

5. [18 pt] Bij huiswerkopgave 3 werd gevraagd om een context-vrije grammatica voor de taal

$$L = \{a^i b^j c^k \mid i, j, k \geq 0 \text{ en } i + k + 2 < j\}$$

De eerste zes elementen in de canonieke volgorde van L zijn bbb, $bbbb$, $abbbb$, $bbbbb$, $bbbbc$, $abbbbb$.

Beschouw nu de volgende drie context-vrije grammatica's G_1, G_2 en G_3:

(a) G_1 heeft startvariabele S en de volgende producties:

$$S \to aSb \mid bbbC \qquad C \to bCc \mid bC \mid \Lambda$$

(b) G_2 heeft startvariabele S en de volgende producties:

$$S \to AbbB \mid AbbBC \mid bbBC \qquad A \to aAb \mid ab \qquad B \to bB \mid b \qquad C \to bCc \mid bc$$

(c) G_3 heeft startvariabele S en de volgende producties:

$$S \to AbbbC \qquad A \to aAb \mid \Lambda \qquad C \to bCc \mid B \qquad B \to bB \mid \Lambda$$

Beantwoord voor elk van deze drie grammatica's G_i de volgende twee vragen:

(i) Is $L \subseteq L(G_i)$? Zo nee, geef een string x die wel in L zit, maar niet in $L(G_i)$. Zo ja, dan hoef je dat niet toe te lichten.

(ii) Is $L(G_i) \subseteq L$? Zo nee, geef een string x die wel in $L(G_i)$ zit, maar niet in L. Zo ja, dan hoef je dat niet toe te lichten.

---

6. [12 pt] Laat G_1 de context-vrije grammatica zijn met startvariabele S en de volgende producties:

$$S \to Sa \mid bb \mid AB \qquad A \to aAb \mid BBa \qquad B \to SB \mid a \mid \Lambda$$

In deze opgave gaan we de eerste stappen uitvoeren van een algoritme om G_1 om te zetten in Chomsky normaalvorm.

(a) Bepaal stap voor stap (dus via $N_0, N_1, N_2, \ldots$) de *nullable* variabele(n) in G_1.

(b) Geef de context-vrije grammatica G_2 die ontstaat uit G_1 door Λ-producties te elimineren.

(c) Geef voor elke variabele X in G_2 de verzameling van X-*derivable* variabelen.

(d) Geef de context-vrije grammatica G_3 die ontstaat uit G_2 door unitproducties te elimineren.

Z.O.Z.

7. [11 pt] Laat opnieuw

$$L = \{a^i b^j c^k \quad | \quad i, j, k \geq 0 \text{ en } i + k + 2 < j\}$$

Teken een stapelautomaat M, zó dat $L(M) = L$.

Deze stapelautomaat moet rechtstreeks gebaseerd zijn op eigenschappen van de taal. Hij moet dus niet het resultaat zijn van een standaardconstructie om bijvoorbeeld een context-vrije grammatica om te zetten in een stapelautomaat.

Probeer ervoor te zorgen dat M deterministisch is en geen Λ-transities bevat. Als dat niet lukt, kun je nog wel een deel van de punten verdienen.

Leg ook uit hoe M zijn toestanden en stapel gebruikt om precies de juiste taal te accepteren.

---

8. [13 pt] Het pomplemma voor context-vrije talen luidt als volgt:

> Stel L is een context-vrije taal.
> Dan is er een integer $n \geq 2$, zó dat
> voor iedere $u \in L$ waarvoor $|u| \geq n$, u geschreven kan worden als
> $u = vwxyz$ voor bepaalde strings v, w, x, y en z waarvoor
> 1. $|wy| \geq 1$ (dwz $wy \neq \Lambda$).
> 2. $|wxy| \leq n$.
> 3. Voor elke $m \geq 0$ behoort de string $vw^m xy^m z$ ook tot L.

Laat nu

$$L_1 = \{x \in \{a, b, c\}^* \quad | \quad n_a(x) = n_c(x) \text{ en } n_a(x) + n_c(x) \leq n_b(x) + 1\}$$

Dus bijvoorbeeld $a^{10}c^{10}b^{19} \in L_1$. Laat nu n voor deze taal het getal uit het pomplemma zijn. Je mag ervan uitgaan dat $n \geq 4$.

Geef bij elk van de onderstaande vier strings u_1, u_2, u_3, u_4 aan of hij geschikt is om een tegenspraak met het pomplemma te vinden. Geef bovendien voor elke string u_i die **niet** geschikt is, aan waarom hij niet geschikt is, bijvoorbeeld via een concrete opsplitsing $vwxyz$ van u_i die keurig aan het pomplemma voldoet. Als u_i wel geschikt is om een tegenspraak met het pomplemma te vinden, hoef je dat niet toe te lichten.

$$u_1 = a^n b^{2n-2} c^n$$
$$u_2 = a^n b^{2n-1} c^n$$
$$u_3 = a^n b^{3n} c^n$$
$$u_4 = (abbc)^n abc$$

`einde tentamen`