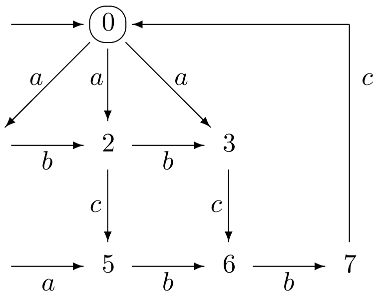
Uitwerking Talen en Automaten, 23 augustus 2006

Tijdsduur 3 uur. Gesloten boek tentamen.

Opgave 1 (20 %). Beschouw de nondeterministische eindige automaat A over het alfabet $\Sigma = \{a, b, c, d\}$ gegeven in de figuur:

A :



De toestanden zijn genummerd 0 tot en met 7. Toestand 0 is de starttoestand en de enige accepterende toestand.
(a) Zet de automaat A volgens het standaardalgoritme om in een deterministische eindige automaat B die dezelfde taal accepteert. Hoeveel bereikbare toestanden heeft B?
(b) Bepaal (bv. op grond van het resultaat van (a)) een reguliere expressie voor de taal van de automaten A en B.

Uitwerking. (a) We maken volgens het standaardalgoritme een tabel van de toestanden van de deterministische automaat B met zijn overgangen:

	δ	a	b	c
$\to$	$\{0\}$	$\{1,2,3\}$	$\emptyset$	$\emptyset$
	$\{1,2,3\}$	$\emptyset$	$\{2,3\}$	$\{4,5,6\}$
	$\{2,3\}$	$\emptyset$	$\{3\}$	$\{5,6\}$
	$\{3\}$	$\emptyset$	$\emptyset$	$\{6\}$
	$\{4,5,6\}$	$\{5\}$	$\{6,7\}$	$\emptyset$
	$\{5,6\}$	$\emptyset$	$\{6,7\}$	$\emptyset$
	$\{5\}$	$\emptyset$	$\{6\}$	$\emptyset$
	$\{6\}$	$\emptyset$	$\{7\}$	$\emptyset$
	$\{6,7\}$	$\emptyset$	$\{7\}$	$\{0\}$
	$\{7\}$	$\emptyset$	$\emptyset$	$\{0\}$

Het symbool d geeft geen overgangen, het heeft daarom geen kolom in de tabel gekregen. De lege toestand $\emptyset$ heeft geen overgangen, $\emptyset$ heeft daarom geen rij in de tabel gekregen. De starttoestand $\{0\}$ is tevens de enige accepterende toestand. Het aantal bereikbare toestanden van B is 11 (inclusief $\emptyset$).

(b) Het is het eenvoudigst om uit te gaan van de automaat A. Om vanuit toestand 0 in 0 te komen, moeten we nul of meer keren vanuit 0 vertrekken en dan via 6 en 7 terugkomen. De gevraagde reguliere expressie E is dus van de vorm $E = (Fbc)^*$ waarbij F de reguliere expressie is van de taal die hoort bij de eindige automaat A' die uit A ontstaat door toestand 7 met zijn overgangen weg te laten en alleen toestand 6 als accepterend te nemen. Er zijn voor F diverse uitdrukkingen mogelijk. Eén ervan is $F = acab + (ab + a)(cb + bc) + ac$, waarbij de term $acab$ hoort bij het pad via 4, en ac hoort bij het pad met de overgang $0 \to 3$, terwijl de middelste term correspondeert met de vier paden via 2.

Opgave 2 (16 %). Beschouw de taal L_2 over het alfabet $\Sigma = \{0, 1, 2\}$, die bestaat uit alle strings met oneven lengte waarvan het middelste symbool een 2 is.
(a) Laat zien dat de taal L_2 contextvrij is.

(b) Formuleer het Pomplemma voor *reguliere* talen.
(c) Bewijs dat de taal L_2 niet regulier is.

Uitwerking. (a) Als de string lengte 1 heeft, is het de string "2". Als de lengte > 1 is, blijft de string in de taal door het eerste en laatste symbool er af te halen. De taal L_2 wordt dus voortgebrcht door de contextvrije grammatica

$$\begin{aligned} &S \to 2 \mid ASA \\ &A \to 0 \mid 1 \mid 2 \ . \end{aligned}$$

met het startsymbool S. Dus L_2 is contextvrij.

(b) Het Pomplemma voor reguliere talen luidt als volgt. Voor elke reguliere taal L is er een pompconstante $n \in \mathbb{N}$ zodanig dat voor elke $w \in L$ met $|w| \geq n$ er strings x, y en z zijn met $xyz = w$ en $|xy| \leq n$ en $|y| > 0$ en met voor alle $k \in \mathbb{N}$ de eigenschap $xy^kz \in L$.

(c) Stel dat L_2 regulier is. Laat n zijn pompconstante zijn. Beschouw de string $w = 0^n21^n$. Hiervoor geldt $w \in L_2$ en $|w| \geq n$. Omdat n een pompconstante voor L is, zijn er dus strings x, y en z zijn met $xyz = w$ en $|xy| \leq n$ en $|y| > 0$ en met voor alle $k \in \mathbb{N}$ de eigenschap $xy^kz \in L_2$. Uit $xyz = w = 0^n21^n$ en $|xy| \leq n$ volgt dat xy uitsluitend nullen bevat. Stel $|y| = m$, dan is $m > 0$ en dus is $xy^2z = 0^{n+m}21^n \notin L_2$ omdat de enige 2 niet in het midden zit. Dit is een tegenspraak. Dus L_2 is niet regulier.

Opgave 3 (20 %). Beschouw de stapelautomaat $P = (Q, \Sigma, \Gamma, \delta, q_0, Z)$ met $Q = \{q_0, q_1, q_2\}$ en $\Sigma = \{a, b\}$ en $\Gamma = \{Z\}$. De functie δ voldoet aan

$$\begin{aligned} &\delta(q_0, a, Z) = \{(q_0, ZZZ)\} \ \text{ en } \ \delta(q_0, b, Z) = \{(q_1, Z)\} \ , \\ &\delta(q_1, a, Z) = \{(q_1, \varepsilon)\} \ \text{ en } \ \delta(q_1, b, Z) = \delta(q_2, \varepsilon, Z) = \{(q_2, \varepsilon)\}. \end{aligned}$$

Alle andere waarden van $\delta(q, u, Z)$ voor $q \in Q$ en $u \in \Sigma \cup \{\varepsilon\}$ zijn leeg. L_3 is de taal geaccepteerd door de stapelautomaat P bij lege stapel.
(a) Beschrijf de taal L_3 volledig als een verzameling van strings. Hoeveel symbolen a en b kunnen deze strings bevatten, en in welke volgorde?
(b) Beschrijf een andere stapelautomaat die dezelfde taal L_3 accepteert, maar die per stap nooit meer dan twee symbolen op de stapel zet.
(c) Geef een contextvrije grammatica die de taal L_3 voortbrengt (hetzij door middel van het standaardalgoritme, hetzij op grond van je beschrijving in onderdeel (a)).

Uitwerking. (a) De stapelautomaat P heeft alleen overgangen $q_0 \to q_1$ en $q_1 \to q_2$. In q_0 wordt de stapel alleen groter. De toestand blijft alleen q_0 door het accepteren van a's. Toestand q_0 wordt alleen verlaten door een b te accepteren. Elke string van L_3 begint dus met a^nb voor zekere $n \in \mathbb{N}$; dan is de automaat in toestand q_1 gekomen en de stapel bevat Z^{2n+1}.

De stapel kan nu op twee manieren leeg raken: hetzij door in q_1 te blijven en $2n+1$ symbolen a te accepteren, hetzij door eerst $k \leq 2n$ symbolen a te accepteren en dan met een b naar toestand q_2 te gaan, alwaar de stapel zonder verdere invoer wordt afgebouwd. Dit levert dus de taal

$$L_3 = \{a^nba^{2n+1} \mid n \in \mathbb{N}\} \cup \{a^nba^kb \mid k, n \in \mathbb{N} : k \leq 2n\} \ .$$

(b) Er zijn diverse oplossingen. Eén ervan gaat alsvolgt. De stapelautomaat P' krijgt een extra stapelsymbool X en de overgangsfunctie δ' met $\delta'(q_0, a, Z) = \{(q_0, XZ)\}$ en $\delta'(q_0, \varepsilon, X) = \{(q_0, ZZ)\}$. Voor alle andere parameterwaarden is $\delta' = \delta$. Er worden door P' nooit meer dan twee symbolen tegelijk op de stapel gezet. Als er een X op de stapel gezet wordt, moet deze X de volgende stap door ZZ vervangen worden. Daarom heeft deze automaat dezelfde taal als P.

(c) De strings uit de taal van onderdeel (a) eindigen op een a of een b. Wat daaraan vooraf gaat, wordt voortgebracht door de respectievelijke nonterminals A en B volgens de contextvrije grammatica

$$S \to Aa \mid Bb$$
$$A \to b \mid aAaa$$
$$B \to b \mid aB \mid aBa \mid aBaa \; .$$

met startsymbool S. Dit is dus een grammatica voor L_3.

Opgave 4 (14 %). L_4 is de taal over het alfabet $\Sigma = \{a, b, c\}$ die bestaat uit alle strings die evenveel symbolen a als b bevatten. Beschrijf een eenbands Turing machine M die de taal L_4 accepteert; geef het volledige zeven-tupel.

Uitwerking. Er zijn weer diverse oplossingen. De mooiste gevonden oplossing gaat aldus. Gebruik het bandalfabet $\Gamma = \Sigma \cup \{X, B\}$ en het volgende algoritme:

$q0$: zoek met de kop naar rechts naar het eerste symbool a of b, vervang het door X, en ga naar $q1$ als het een a was en naar $q2$ als het een b was. Als een B gevonden wordt, ga naar $q4$.

$q1$: zoek met de kop naar rechts het eerste symbool b, vervang het door c en ga naar $q3$ en links; indien afwezig, verwerp de invoer;

$q2$: zoek met de kop naar rechts het eerste symbool a, vervang het door c en ga naar $q3$ en links; indien afwezig, verwerp de invoer;

$q3$: zoek met de kop naar links het symbool X, ga naar rechts en ga naar $q0$;

$q4$: accepteer de invoer.

Dit is rechtstreeks om te zetten in een Turing machine $M = (Q, \Sigma, \Gamma, \delta, q0, B, F)$ met $Q = \{q0, \ldots, q4\}$, en Σ en Γ als boven, accepterende verzameling $F = \{q4\}$, en de overgangsfunctie δ volgens

δ	a	b	c	X	B
$q0$	$(q1, X, R)$	$(q2, X, R)$	$(q0, c, R)$	?	$(q4, ?, ?)$
$q1$	$(q1, a, R)$	$(q3, c, L)$	$(q1, c, R)$	?	– (verwerp)
$q2$	$(q3, c, L)$	$(q2, b, R)$	$(q2, c, R)$	?	– (verwerp)
$q3$	$(q3, a, L)$	$(q3, b, L)$	$(q3, c, L)$	$(q0, ?, R)$	?

De waarden van de vraagtekens is niet van belang, evenmin als de overgangen vanuit de accepterende toestand $q4$.

Opgave 5 (12 %). We beschouwen talen over een alfabet Σ en Turing machines met invoeralfabet Σ.

(a) Wanneer is een taal L *recursief opsombaar*? Geef de definitie.

(b) Wanneer is een taal L *beslisbaar* (oftewel recursief)? Geef de definitie.

Gegeven is een voorschrift T, dat voor elke string $w \in \Sigma^*$ een Turing machine $T(w)$ vast legt. Dit voorschrift voldoet aan de eigenschap, dat er voor elke Turing machine M een string w bestaat zodanig dat M and $T(w)$ dezelfde taal accepteren. Beschouw nu de taal $L_5 = \{w \in \Sigma^* \mid w \notin L(T(w))\}$.

(c) Is L_5 recursief opsombaar? Bewijs je bewering.

(d) Is L_5 beslisbaar? Bewijs je bewering.

Uitwerking. (a) en (b) Zie boek.

(c) Het antwoord is nee. Dit wordt alsvolgt bewezen. Stel dat L_5 recursief opsombaar is. Dan is er een Turing machine M over Σ met $L_5 = L(M)$. Volgens het gegeven is er dan een string w met $L(T(w)) = L(M) = L_5$. De string w voldoet dus aan

$$w \in L_5 \quad \equiv \quad w \notin L(T(w)) \quad \equiv \quad w \notin L_5 \; .$$

Dit is een tegenspraak. Dus L_5 is niet recursief opsombaar.

(d) Omdat L_5 niet recursief opsombaar is, is L_5 zeker niet beslisbaar.

Opgave 6 (18 %). Gegeven zijn de talen L_a en L_b, en deterministische eenbands Turing machines M_a en M_b waarvoor geldt $L(M_a) = L_a$ en $L(M_b) = L_b$. Concatenatie geeft de taal $L_6 = L_aL_b$.
(a) Beschrijf een nondeterministische (mogelijk meerbands) Turing machine M die voldoet aan $L(M) = L_6$. Bewijs de laatste gelijkheid.
(b) De constructie van onderdeel (a) bewijst een bewering van de vorm: "Als de talen L_a en L_b ... zijn, dan is de concatenatie L_aL_b ook ...." Welke bewering is dit?

Uitwerking. (a) M heeft als bandalfabet de vereniging van de bandalfabetten van M_a en M_b. We nemen aan dat de toestandsruimten Q_a en Q_b van M_a en M_b disjunct zijn, met starttoestanden q_a en q_b. M heeft als toestandsruimte de vereniging $\{q_0, q_1\} \cup Q_a \cup Q_b$, waarbij q_0 en q_1 nieuw zijn en q_0 de nieuwe starttoestand is. M heeft twee banden. De invoer staat op de eerste band. De tweede band is initieel leeg.

In toestand q_0 kopieert M een nondeterministisch beginstuk x van de invoerstring $w = xy$ naar de tweede band en gaat naar toestand q_1. De kop van de tweede band wordt teruggebracht naar de eerste positie van x, waarna de toestand q_a wordt. Vervolgens simuleert M de TM M_a op de tweede band. Als hij in een accepterende toestand komt, vervolgt M met simulatie van M_b op de eerste band ter acceptatie van y.

Als M een accepterende berekening heeft, geldt $x \in L_a$ en $y \in L_b$ en dus $w \in L_6$. Als $w \in L_aL_b$, dan heeft M een accepterende berekening. Dit bewijst $L = L(M)$.

(b) Dit is de bewering: "Als de talen L_a en L_b recursief opsombaar zijn, dan is de concatenatie L_aL_b ook recursief opsombaar."