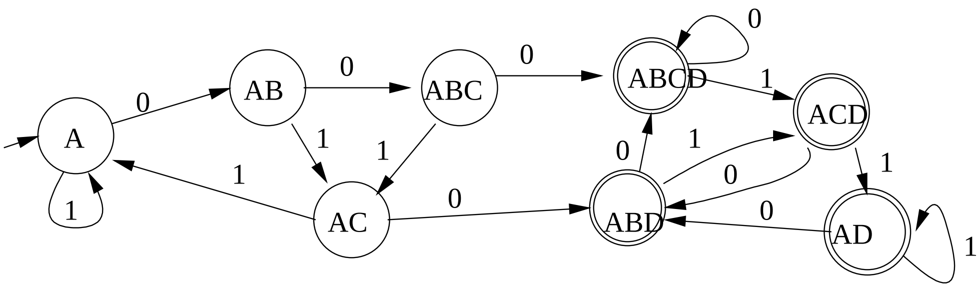
Formális nyelvek vizsga-megoldás

2001. június 12.

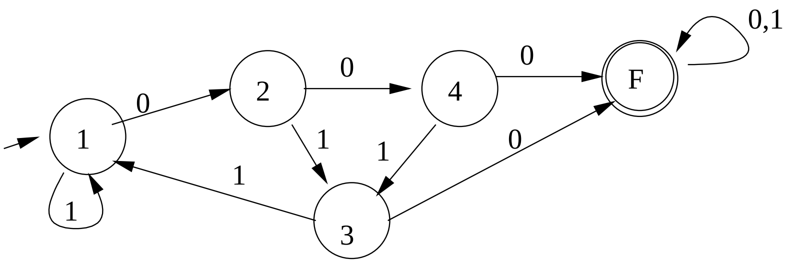
1. Az automata nem determinisztikus, ezért először determinizálni kell:



Ez már egyben teljesen specifikált is, úgyhogy jöhet a minimalizálás:

(A=1, AB=2, AC=3, ABC=4, ABCD=5, ABD=6, ACD=7, AD=8)

1234——5678
12——34——5678
1——2——3——4——5678
És ez már nem változik, tehát a minimálautomata:



2. Igaz, az L nyelv mindig reguláris lesz. Ezt a következőképp lehet megmutatni: Legyen a 37 nyelv $L_1, \ldots, L_{37}$; mivel ezek regulárisak, ezért vannak hozzájuk determinisztikus, teljesen spcifikált véges automaták: $M_1, \ldots M_{37}$. Képezzük ezen automaták direkt szorzatát, azaz futassuk párhuzamosan a 37 automatát. Tehát azt az M véges automatát készítjük el, melynek állapotai 37 komponensből állnak, az i-edik komponens megfelel az M_i automata állapotának. Az M automata kezdőállapota legyen az $S_1 \times \ldots \times S_{37}$ állapot, ahol S_i az M_i automata kezdőállapota. Elfogadó állapot az legyen M-ben, melyben a 37 komponens közül legalább 19 elfogadó a megfelelő M_i véges automatákban. Az átmenet pedig legyen úgy definiálva M-ben, hogy $\delta((q^1, \ldots, q^{37}), a) = (p^1, \ldots, p^{37})$ valami a inputszimbólumra, ha $\delta(q^i, a) = p^i$ az átmenet az M_i automatában, minden i-re.
Így:

$w \in L \Leftrightarrow w$ legalább 19 L_i-ben benne van $\Leftrightarrow w$ hatására legalább 19 M_i elfogadó állapotba jut $\Leftrightarrow$ w hatására M olyan állapotba jut, melynek legalább 19 komponense elfogadó a megfelelő M_i-kben $\Leftrightarrow$ w-t M elfogadja
Tehát az M véges automata az L nyelvet fogadja el.

3. Van hozzá CF nyelvtan:
$S \to aSa \mid aS_1a$, $S_1 \to bS_1b \mid bb$.
Ez a nyelvtan tényleg jó lesz:
1. A generált szavak $a^*b^*a^*$ alakúak lesznek és az a-k száma a szó elején és végén megegyezik; a 2. és a 4. szabály garantálja, hogy $n, m \geq 1$; a b-k szám pedig mindig páros lesz, mert kettesével jönnek be.
2. Minden $a^nb^{2m}a^n$ alakú szó kijön: először $n-1$-szer használjuk az 1. szabályt, aztán egyszer a 2.-t, aztán $m-1$-szer a 3.-t, végül egyszer a 4.-t.
A nyelv nem reguláris, ezért egyszerűbb nyelvtana nincsen. A nem-regularitást a pumpálási lemmával lehet belátni: Ha reguláris lenne, akkor létezne olyan p korlát, hogy minden legalább p hosszú nyelvbeli szó minden legalább p hosszú részszavában lehetne pumpálni valahol. Vegyük ezzel a p-vel az $a^pb^{2p}a^p$ szót és jelöljük ki a kezdő a^p-t. A lemma szerint itt is kell tudni valahol pumpálni. De ezen a részszón belül csak a-sorozatot tudunk kijelölni és ha azt pumpáljuk, akkor a szókezdő a-k száma nem lesz egyenlő a szó végén álló a-k számával, azaz kilépünk a nyelvből. Tehát a nyelv nem reguláris.

4. A nyelvtan: $S \to a_1Sa_2 \mid b_1Sb_2 \mid \epsilon$
A homomorfizmusok: $h_1(a_1) = a$,$h_1(a_2) = \epsilon$,$h_1(b_1) = b$, $h_1(b_2) = \epsilon$,
$h_2(a_1) = a$,$h_2(a_2) = a$,$h_2(b_1) = b$, $h_2(b_2) = b$
Magyarázat: a nyelvtan a $w_1w_2^{-1}$ alakú szavakat generálja, azaz azokat a szavakat, ahol a szó második fele tükörképe a szó első felének, leszámítva, hogy az első félben a betűk indexe 1, a hátsóban pedig 2. Az első homomorfizmus törli a második felét a szónak és törli az 1 indexeket, a második homomorfizmus törli az össze indexet. Így $h_1(w_1w_2^{-1}) = w$, míg $h_2(w_1w_2^{-1}) = ww^{-1}$. Épp ez kellett.

5. Epszilon szabály nincs.
Láncszabály kiküszöbölés után: $S \to aB \mid bS \mid b \mid AS$, $A \to aB \mid bS \mid b$, $B \to AB \mid Ba$, $C \to AS \mid b$
Felesleges szimbólumok kiküszöbölése, először alulról felfele: kiesik a B, azaz: $S \to bS \mid b \mid AS$, $A \to\ bS \mid b$, $C \to AS \mid b$
aztán fentről lefele: kiesik a C, azaz: $S \to bS \mid b \mid AS$, $A \to\ bS \mid b$,
CNF-ra hozva:
$S \to \hat{B}S \mid b \mid AS$, $A \to\ \hat{B}S \mid b$,$\hat{B} \to b$

6. Biztos nem $LL(0)$ elemezhető, mert nem egy szóból áll a generált nyelv. Nézzük az $LL(1)$-et:

	S_1	$\{ \epsilon \}$
a, b, ϵ	A_1B_1,1	$\{b, \epsilon\}, \{\epsilon\}$

	A_1	$\{ b, \epsilon \}$
a	aA_2b,2	$\{b\}$
b, ϵ	ϵ,3	

	B_1	$\{ \epsilon \}$
b	bB_2c,4	$\{ c \}$
ϵ	ϵ,5	

	A_2	$\{ b\}$
a	aA_2b,2	$\{b\}$
b	ϵ,3	

	B_2	$\{ c \}$
b	bB_2c,4	$\{ c \}$
c	ϵ,5	

Tehát elemezhető $k = 1$-re és az elemző tábla:

	a	b	c	ϵ
S_1	A_1B_1,1	A_1B_1,1		A_1B_1,1
A_1	aA_2b,2	ϵ,3		ϵ,3
A_2	aA_2b,2	ϵ,3		
B_1		bB_2c,4		ϵ,5
B_2		bB_2c,4	ϵ,5	
a	pop			
b		pop		
c			pop	
ϵ				A

Van erős elemző is $k = 1$-re, mert a táblázatban sem az A_1 és A_2, sem a B_1 és B_2 sora nem ütközik, össze lehetne húzni őket.

Az elemzés:
$(S_1\ ,\ abbc\ ,\ \epsilon) \rightarrow$
$(A_1B_1\ ,\ abbc\ ,\ 1) \rightarrow$
$(aA_2bB_1\ ,\ abbc\ ,\ 12) \rightarrow$
$(A_2bB_1\ ,\ bbc\ ,\ 12) \rightarrow$
$(bB_1\ ,\ bbc\ ,\ 123) \rightarrow$
$(B_1\ ,\ bc\ ,\ 123) \rightarrow$
$(bB_2c\ ,\ bc\ ,\ 1234) \rightarrow$
$(B_2c\ ,\ c\ ,\ 1234) \rightarrow$
$(c\ ,\ c\ ,\ 12345) \rightarrow$
$(\epsilon\ ,\ \epsilon\ ,\ 12345) \rightarrow$
Accept