7. 10. 2011

Regulární gramatiky

abeceda $T = \{0, 1\}$ chceme, aby přijímal lichý počet 1

gramatika

$G = (\{ \quad \}, \{0, 1\}, P, S)$ (může být libovolný znak)

nonterminály, zatím neznáme (velkým písmenem)

terminály (vstupní symboly) (malým písmenem)

→ startovní symbol

→ pravidla (budou dále)

⇒ zajímá nás sudý počet 1

$S \rightarrow 0$ – nezajímá nás ⇒ na počtu 1 se nic nemění

$S \rightarrow 0S$ → neterminální symbol sudý počet 1 se nezměnil

$P: \rightarrow S \rightarrow 0S \mid 1 \mid 1L$

= máme sudý počet jedniček

→ pokud přijde 1, je počet 1 lichý, což je věta daného jazyka ⇒ můžeme nechat ~~tak~~ 1 samostatně

→ sudý počet 1

→ pokud chceme, aby za 1 bylo něco dalšího – dáme 1 a dostáváme tím lichý počet 1 = L

$L \rightarrow 0L \mid 1S \mid 0$

↳ máme lichý počet 1

↳ počet 1 se nemění => zůstává lichý

↳ pokud přijde 1, stává se počet 1 sudým

↳ pokud by přišla 0, pak je to OK, protože počet 1 zůstává lichý => přijímáme

doplníme neterminální symboly - stačí dva - S, L

$G_1 = (\{S, L\}, \{0, 1\}, P, S)$

$S \rightarrow 0S \mid 1 \mid 1L$

$L \rightarrow 0L \mid 1S \mid 0$

převedeme gramatiku na konečný automat

$KA = (\{S, L, K\}, \{0, 1\}, \delta, S, \{K\})$

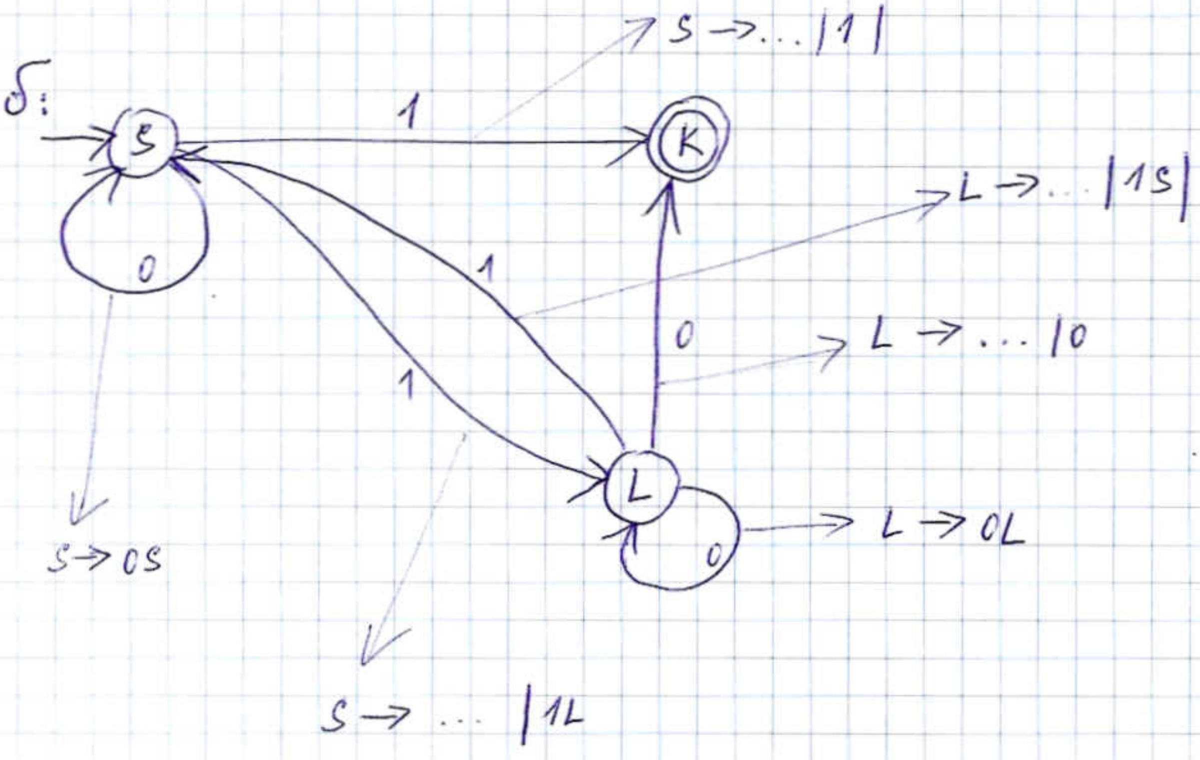
↳ vstupní abeceda

↳ přechodová relace

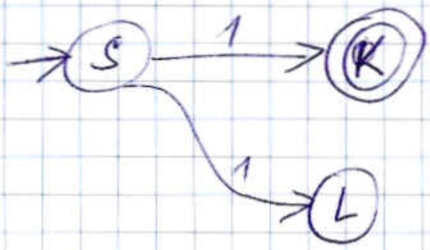
↳ množina stavového automatu (zatím neznáme - doplníme postupně)

↳ počáteční stav

↳ množina koncových stavů, tady bude jediný



Vytvořený KA <u>není</u> deterministický ← máme dva přechody pro jeden a ten samý vstupní symbol, tj. 1



PŘ $T = \{0,1\}$

sudý počet 0, počet 1 dělitelný 3 beze zbytku ⟸ navrhnout takovou gramatiku

$G = (\{S_0, S_1, L_0, L_1, S_2, L_2\}, \{0,1\}, P, S_0)$

$P: \rightarrow S_0 \rightarrow 0L_0 \mid 1S_1 \mid \varepsilon$

↳ ~~počet 1~~ když zbytek po dělení 1 = 0

↳ pokud přijde 1, zbytek po dělení = 1

↳ prázdný řetěz = řetěz 0 terminálních symbolů

$L_0 \rightarrow 0S_0 \mid 1L_1$

gramatika je lineární

$S_1 \rightarrow 0L_1 \mid 1S_2$

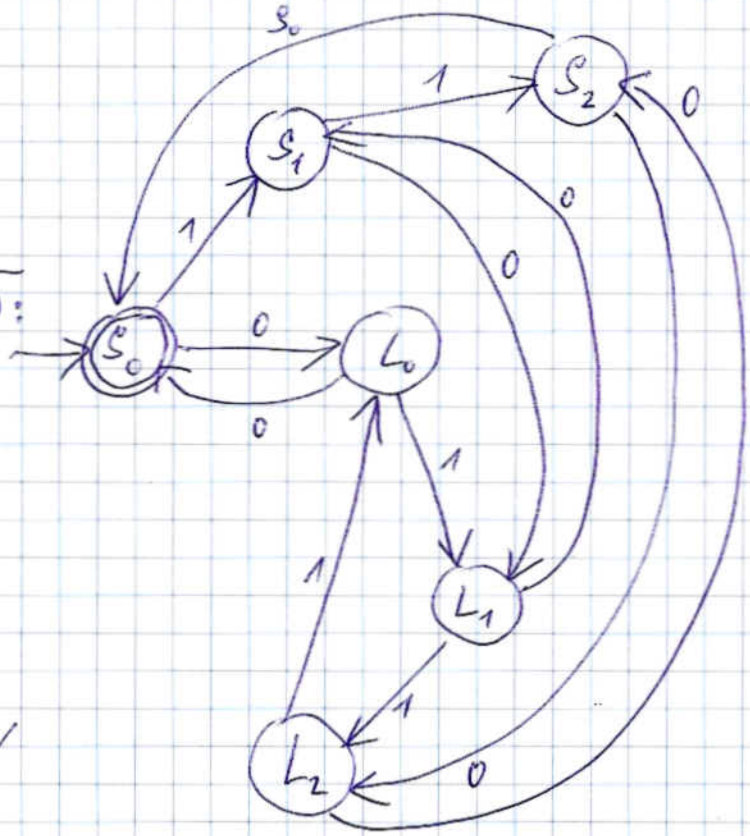
$L_1 \rightarrow 0S_1 \mid 1L_2$

$S_2 \rightarrow 0L_2 \mid 1S_0$

$L_2 \rightarrow 0S_2 \mid 1L_0$

převedení na KA

δ:



$KA = (\{S_0, L_0, S_1, L_1, S_2, L_2\}, \{0,1\}, \delta, S_0, \{S_0\})$

KA je deterministický

PŘ $LG \to RG$ převod lineární gram. na regulární

$\to S \to 0A \mid 1B \mid 1S \mid \varepsilon$

$A \to 101S \mid B$

$B \to 0S \mid 1A \mid \varepsilon$

RG musí splňovat pravidla - viz přednášky

co je ~~v~~ nutno odstranit
- ε v S symbolu
- 101S (= 3 terminální + 1 neterminální symbol) v A

$A \to 1C \mid \underbrace{0S \mid 1A \mid \varepsilon}$

↳ zavedení nového symbolu

↳ dosazení místo B, což je druhá možnost, která může přijít

C → nutno předělat, když jsme ho vytvořili

~~[illegible]~~ v $A \to 10\boxed{1S} \mid B$ → D

↳ C

$C \to 0D$

$D \to 1S$

stav v $S \to 0A \mid 1B \mid 1S \mid \varepsilon$

~~$A \to 101S \mid B$~~

$B \to 0S \mid 1A \mid \varepsilon$

$A \to 1C \mid 0S \mid 1A \mid \varepsilon$

$C \to 0D$

$D \to 1S$

nyní odstraníme ε pravidla

<u>obecně</u> - neterminální symbol ~~??~~ přepiš na ε, pak musíme najít daný neterminální symbol na pravých stranách všech pravidel a ε za něj dosadit

- pokud máme ε pravidlo jako jediné, pak se dá neterminální symbol ε přímo nahradit
- pokud ~~x~~ máme více pravidel a jedno z nich je ε, tak musíme na všech P stranách, kde se neterminální symbol vyskytuje, přidat další pravidlo, kde terminální symbol zůstane a neterminální se nahradí ε

rušíme ε

→ B přepiš na ε

$S \rightarrow 0A \mid 1B \mid 1S \mid \cancel{\varepsilon} \mid 0 \mid 1$

~~$A \rightarrow 101S \mid B$~~ → A přepiš na ε

$B \rightarrow 0S \mid 1A \mid \cancel{\varepsilon} \mid 1 \mid 0$

$A \rightarrow 1C \mid 0S \mid 1A \mid \cancel{\varepsilon} \mid 1 \mid 0$

$C \rightarrow 0D$

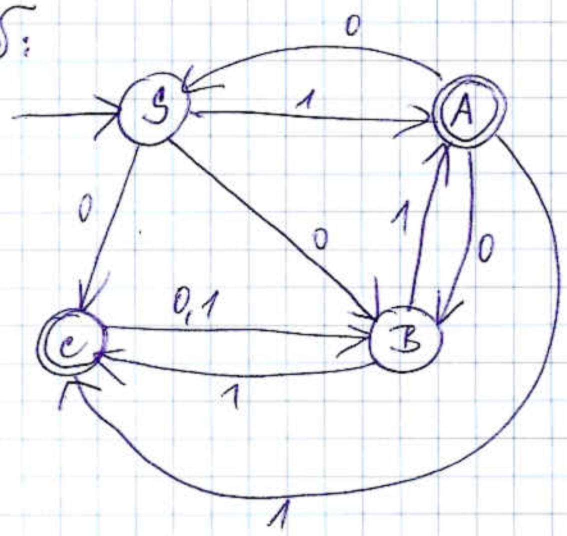
$D \rightarrow 1S$

$S_0 \rightarrow 0A \mid 1B \mid 1S \mid \varepsilon \mid 0 \mid 1$

S přepiš na ε - uděláme tak, že zavedeme nový start. symbol, např. S_0 pak lze odstranit ε z S

↳ zde není nutno rušit, protože S_0 není na pr. straně žádného pravidla

(PŘ) δ:



převod KA determinist. na nedeterministický

$\{ \dots \}$ = množina
$[\dots]$ = stav

$G = (\{S, A, B, C\}, \{0,1\}, \delta, S, \{A, C\})$

přepis do tabulky

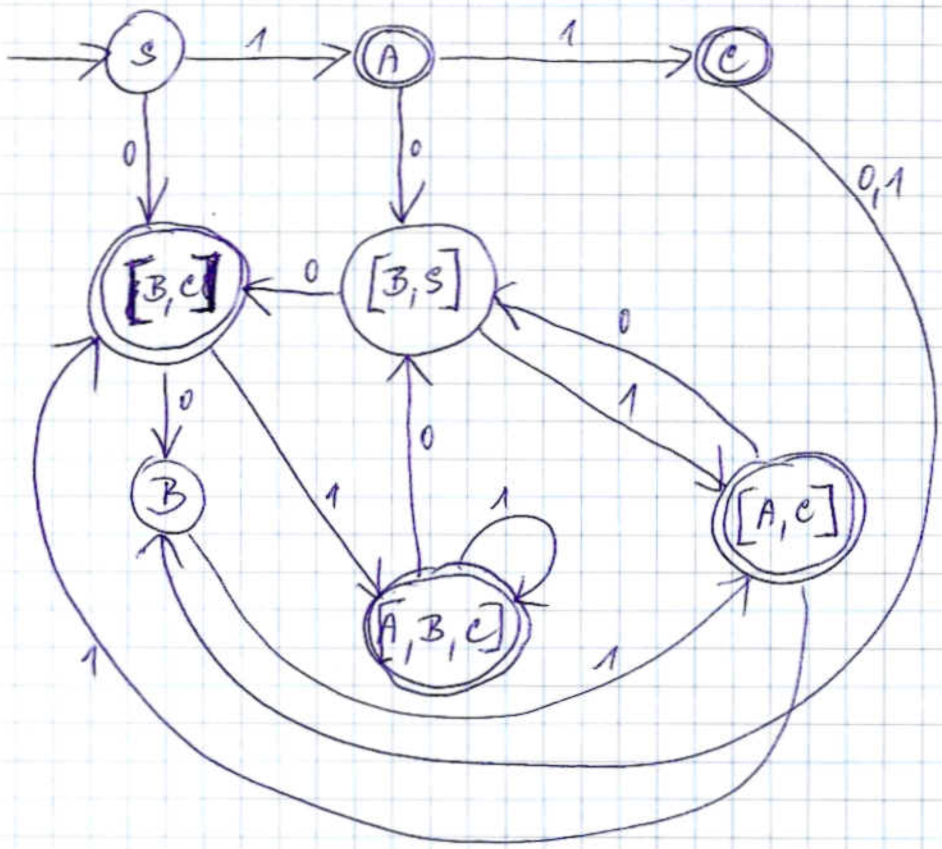
δ	0	1
$\rightarrow S$	B, C	A
$\leftarrow A$	S, B	C
B		A, C
$\leftarrow C$	B	B

	0	1
$\rightarrow S$	[B, C]	A
$\leftarrow A$	[B, S]	C
$\leftarrow$ [B, C]	B	[A, B, C]
[S, B]	[B, C]	[A, C]
$\leftarrow C$	B	B
B		[A, C]
$\leftarrow$ [A, B, C]	[S, B]	[A, B, C]
$\leftarrow$ [A, C]	[S, B]	[B, C]

neznačí se, ale můžeme si označit

převod – u KA jde VŽDY !!!
– budeme procházet všechny možnosti paralelně

max. počet vnitřních stavů – 2^n



koncové stavy – všechny, které obsahují některé z původních koncových stavů

PŘ

$T = \{0, 1\}$, obsahuje ~~1~~ 1011

nadefinujte gramatiku, která obsahuje všechny řetězy takové, které obsahují sekvenci 1011 obsaženou kdekoli

~~#~~ předpoklad - máme libovolný počet 0 a 1.

$\rightarrow S \rightarrow \underbrace{0S \mid 1S}_{} \mid 1A$

↳ první znak sekvence

$A \rightarrow 0B$

$B \rightarrow 1C$

$C \rightarrow 1D$

$D \rightarrow \varepsilon \mid \underbrace{0D \mid 1D}_{}$

↳ jsme na konci sekvence

↳ nebo generujeme ~~nějaký~~ ~~nezajímavý~~ nějaký nezajímavý prefix

gramatika je lineární => nutno převést na regulární

↳ na P straně je u D ε

$\rightarrow S \rightarrow 0S \mid 1S \mid 1A$

$A \rightarrow 0B$

$B \rightarrow 1C$

$C \rightarrow 1D \mid 1$

$D \rightarrow \cancel{\varepsilon} \mid 0D \mid 1D \mid 0 \mid 1$

přepis do tabulky

	0	1
S	S	S, A
A	B	
B		C
C		D, K
D	D, K	D, K
← K		

nový pomocný stav

	0	1
→ S	S	S, A
A	B	
B		C
C		D, K
D	D, K	D, K
← K		

	0	1	
→ S	S	[S, A]	1.
[S, A]	[S, B]	[S, A]	2.
[S, B]	S	[S, A, C]	3.
[S, A, C]	[S, B]	[S, A, D, K]	4.
← [S, A, D, K]	[S, B, D, K]	[S, A, D, K]	5.
← [S, B, D, K]	[S, D, K]	[S, A, C, D, K]	6.
← [S, D, K]	[S, D, K]	[S, A, D, K]	7.
← [S, A, C, D, K]	[S, A, D, K]	[S, A, D, K]	8.

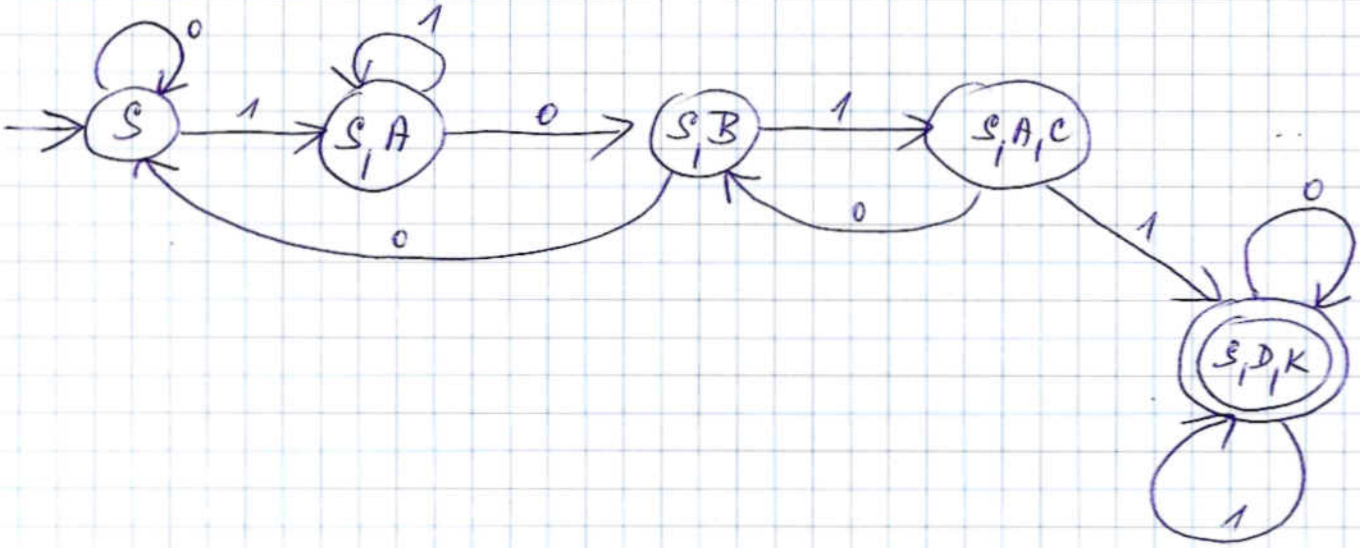
Provedeme minimalizaci stavů

- hledáme stavy, které mají tzv. „stejné řádky" (bez záhlaví)
- lze sloučit dva → nebo ← stavy, a bez →/←, nelze sloučit → a ← stav, ~~nebo bez →/←~~ ~~dva~~

Lze sloučit ř. 5. a 8. ← shoda v {1}

Provedení – škrtneme a nahradíme

	0	1	ř.	0	1
→S	S	[S,A]	1.		
[S,A]	[S,B]	[S,A]	2.		
[S,B]	S	[S,A,C]	3.		
[S,A,C]	[S,B]	~~[S,A,D,K]~~	4.		[S,D,K]
~~←[S,A,D,K]~~	~~[S,D,K]~~ ~~[S,B,D,K]~~	~~[S,A,D,K]~~	~~5.~~	zbyde mi [S,D,K]	
~~←[S,B,D,K]~~	~~[S,D,K]~~	~~[S,A,C,D,K]~~	6.		~~[S,A,D,K]~~ +
← [S,D,K]	[S,D,K]	~~[S,A,D,K]~~	7.		~~[S,A,D,K]~~ + [S,D,K]
~~←[S,A,C,D,K]~~	~~[S,A,D,K]~~	~~[S,A,D,K]~~	~~8.~~	zbyde mi [S,A,D,K]	


S
S, A
S, B
S, A, C
S, D, K
0
1