REPREZENTACE JAZYKŮ TYPU 3 POMOCÍ REGULÁRNÍCH VÝRAZŮ

OBDOBA „ARITMETICKÝCH VÝRAZŮ" EFEKTIVNÍ ZPŮSOB POPISU JAZYKA.

REGULÁRNÍ MNOŽINOU NAD ABECEDOU Σ ROZUMÍME KAŽDOU TAKOVOU MNOŽINU ŘETĚZCŮ, KE KTERÉ EXISTUJE KONEČNÝ AUTOMAT, JENŽ JI ROZPOZNÁVÁ.

REGULÁRNÍ MNOŽINY NAD Σ LZE DEFINOVAT REKURZIVNĚ:

1) $\emptyset$ JE RM NAD Σ
2) $\{e\}$ JE RM NAD Σ
3) $\{a\}$ JE RM NAD Σ $\forall a \in \Sigma$
4) JSOU-LI P a Q RM NAD Σ, PAK
   a) $P \cup Q$ JE RM NAD Σ
   b) PQ JE RM NAD Σ
   c) P^* Q^* JSOU RM NAD Σ

5) NEEXISTUJÍ ŽÁDNÉ JINÉ RM NAD Σ.

JINAK ŘEČENO: KAŽDOU RM LZE VYTVOŘIT Z ELEMENTÁRNÍCH RM (1, 2, 3) KONEČNÝM POČTEM APLIKACÍ PRAVIDEL (4a, 4b, 4c).

JAK VYPADAJÍ KA AKCEPTUJÍCÍ ELEMENTÁRNÍ RM?

 $L(A) = \emptyset$

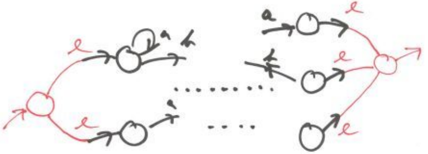
 $L(A) = \{e\}$

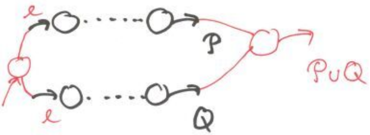
$L(A) = \{a\}$

!POZOR! $\emptyset \neq \{e\}$

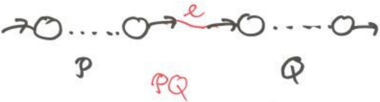
JAK KONSTRUOVAT KA AKCEPTUJÍCÍ RM VZNIKLÉ OPERACEMI ad $\{a, b, c$?

BEZ ÚJMY NA OBECNOSTI PŘEDPOKLÁDEJME, ŽE KAŽDÝ NKA MÁ PRÁVĚ JEDEN VSTUPNÍ A PRÁVĚ JEDEN VÝSTUPNÍ STAV:

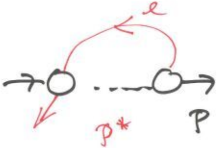




PŘECHODOVÝ GRAF REPREZENTUJÍCÍ AUTOMAT, KTERÝ AKCEPTUJE SJEDNOCENÍ RM.



PŘECHODOVÝ GRAF REPREZENTUJÍCÍ AUTOMAT, KTERÝ AKCEPTUJE ZŘETĚZENÍ RM.



PŘECHODOVÝ GRAF, KTERÝ AKCEPTUJE ITERACI RM.

REGULÁRNÍ VÝRAZY SE POUŽÍVAJÍ K REPREZENTACI RM.

(JEDNODUŠŠÍ ZÁPIS)

REGULÁRNÍ VÝRAZY NAD Σ LZE DEFINOVAT REKURZIVNĚ:

1) $\emptyset$ JE RV NAD Σ OZNAČUJÍCÍ $\emptyset$
2) e JE RV NAD Σ OZNAČUJÍCÍ $\{e\}$
3) a JE RV NAD Σ OZNAČUJÍCÍ $\{a\}$
4) JSOU-LI p A q RV NAD Σ OZNAČUJÍCÍ P A Q, PAK
   a) $p+q$ JE RV NAD Σ OZNAČUJÍCÍ $P \cup Q$
   b) pq JE RV NAD Σ OZNAČUJÍCÍ PQ
   c) p^* A q^* JSOU RV NAD Σ OZNAČUJÍCÍ P^* A Q^*
5) NEEXISTUJÍ ŽÁDNÉ JINÉ RV.

PŘÍKLADY REGULÁRNÍCH VÝRAZŮ
A REGULÁRNÍCH MNOŽIN NAD $\Sigma = \{a, b\}$

RV	RM
ba^*	VŠECHNA SLOVA NAD $\{a, b\}$, ZAČÍNAJÍCÍ PÍSMENEM b, NÁSLEDOVANÝM POUZE ŘETĚZCEM PÍSMEN $\{a\}$ (I PRÁZDNÝM)
$a^*ba^*ba^*$	VŠECHNA SLOVA NAD $\{a, b\}$ OBSAHUJÍCÍ PRÁVĚ DVĚ PÍSMENA b
$(a+b)^*$	VŠECHNA SLOVA NAD $\{a, b\}$
$(a+b)^*aa$	VŠECHNA SLOVA NAD $\{a, b\}$ KONČÍCÍ DVĚMA ZNAKY a.

JAK K RV SESTROJIT NKA AKCEPTUJÍCÍ JAZYK POPSANÝ RV?

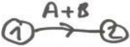
ROZKLADEM ZOBECNĚNÉHO PŘECHODOVÉHO GRAFU.

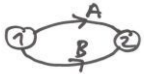
VÝCHOZÍ:

RV POSTUPNĚ ROZKLÁDÁME A PŘIDÁVÁME NOVÉ UZLY A HRANY TAK DLOUHO, DOKUD NENÍ KAŽDÁ HRANA OHODNOCENA PÍSMENEM ZE Σ NEBO SYMBOLEM e.

SCHÉMATA PRO ROZKLAD:

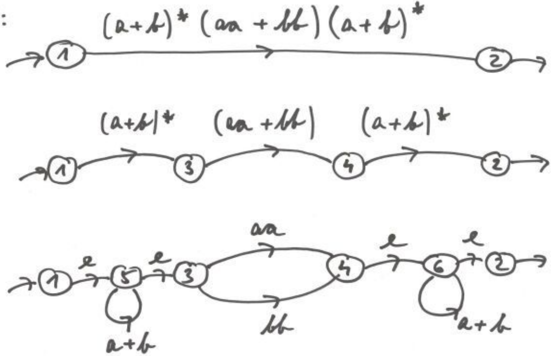
NAHRADÍME

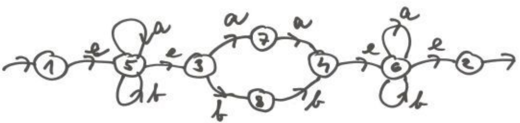



NAHRADÍME


NAHRADÍME


Př:





KAŽDÁ HRANA JE OZNAČENA PÍSMENEM NEBO SYMBOLEM e, ROZKLAD JE TEDY UKONČEN.

AUTOMAT JE OBECNĚ NEDETERMINISTICKÝ S e-HRANAMI. LZE NALÉZT DETERMINISTICKÝ EKVIVALENT. POZDĚJI.

JAK K PŘECHODOVÉMU GRAFU A SESTROJÍME ODPOVÍDAJÍCÍ REGULÁRNÍ VÝRAZ?

[PŘEDPOKLÁDÁME BEZ ÚJMY NA OBECNOSTI NKA S JEDNÍM VSTUPNÍM A JEDNÍM VÝSTUPNÍM STAVEM]

POSTUPNĚ BUDEME ELIMINOVAT VŠECHNY STAVY PŘECHODOVÉHO GRAFU KROMĚ POČÁTEČNÍHO A KONCOVÉHO STAVU. BĚHEM PROCESU ELIMINACE BUDOU MODIFIKOVÁNY HRANY A JEJICH OHODNOCENÍ PŘECHODOVÝMI VÝRAZY. TÍMTO ZPŮSOBEM POSTUPNĚ KONSTRUUJEME ZOBECNĚNÉ PŘECHODOVÉ GRAFY, KTERÉ AKCEPTUJÍ STEJNOU MNOŽINU SLOV. CELÝ PROCES KONČÍ, KDYŽ ZÍSKÁME PŘECHODOVÝ GRAF SE DVĚMA STAVY.

SCHÉMA ELIMINACE HRAN:

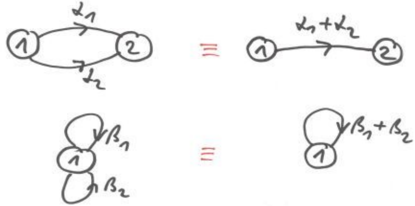
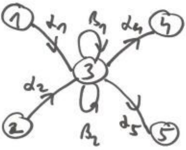


SCHÉMA ELIMINACE VRCHOLŮ:



EL. 3:

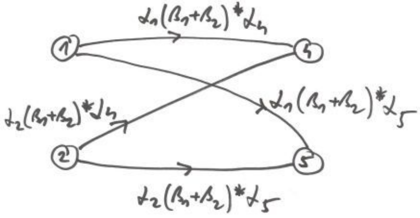
$1 \to 4 \quad \alpha_1(\beta_1+\beta_2)^*\alpha_4$

$1 \to 5 \quad \alpha_1(\beta_1+\beta_2)^*\alpha_5$

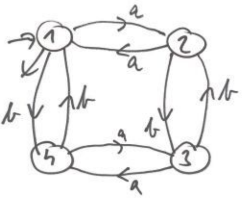
$2 \to 4 \quad \alpha_2(\beta_1+\beta_2)^*\alpha_4$

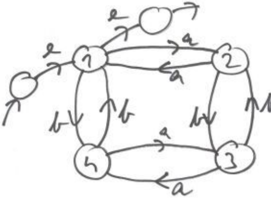
$2 \to 5 \quad \alpha_2(\beta_1+\beta_2)^*\alpha_5$

PO ELIMINACI:



PŘ: AUTOMAT, KTERÝ AKCEPTUJE MNOŽINU VŠECH SLOV NAD {a, b} SE SUDÝM POČTEM {a} A SUDÝM POČTEM {b}.





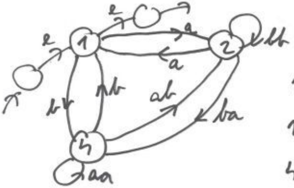
El 3:

2→2 bb

2→4 ba

4→2 ab

4→4 aa



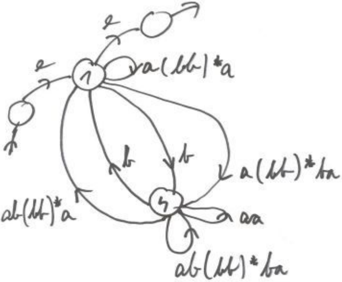
El 2:

1→1 $a(bb)^*a$

1→4 $a(bb)^*ba$

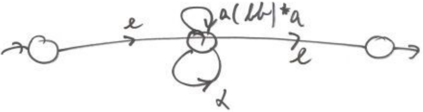
4→1 $ab(bb)^*a$

4→4 $ab(bb)^*ba$



El 4:

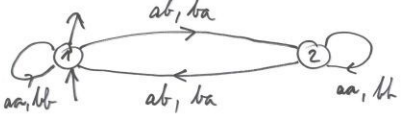
$$1 \to 1 \quad \left[(b+a(bb)^*ba)(aa+ab(bb)^*ba)^* (b+ab(bb)^*a)\right] = \alpha$$

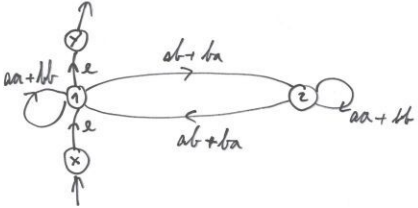


UITSLOEDEN:

$$\rightarrow \bigcirc \xrightarrow{\left[a(bb)^*a + (b+a(bb)^*(ba))(aa+ab(bb)^*ba)^*(b+ab(bb)^*a)\right]^*} \bigcirc \rightarrow$$

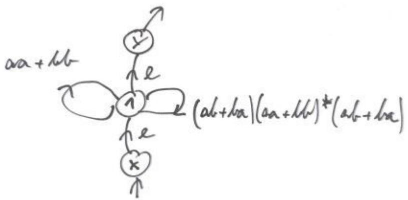
JEDNODUŠŠÍ ŘEŠENÍ





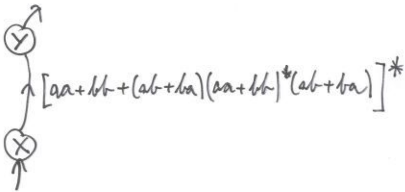
EL 2:

$$1 \to 1 \quad (ab+ba)(aa+bb)^*(ab+ba)$$



El 1:

$$X \to Y : [aa+bb+(ab+ba)(aa+bb)^*(ab+ba)]$$



ZÁVĚR K PŘÍKLADU :

OBA ZÍSKANÉ RV ZŘEJMĚ REPREZENTUJÍ JEDNU REGULÁRNÍ MNOŽINU.

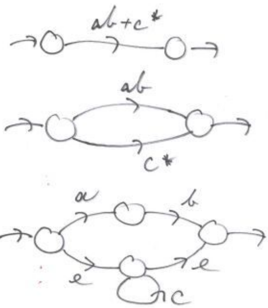
REGULÁRNÍ VÝRAZY JSOU VELICE VARIABILNÍ A NENÍ JEDNODUCHÉ DOKÁZAT, ŽE DVA RŮZNÉ RV JSOU EKVIVALENTNÍ.

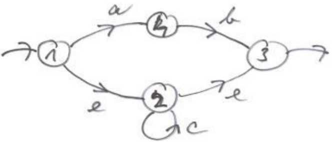
S RV LZE PROVÁDĚT SYMBOLICKÉ ÚPRAVY PODOBNÉ JAKO S ARITMETICKÝMI VÝRAZY (NAPŘ. VYTÝKÁNÍ, ROZNÁSOBOVÁNÍ).

V DALŠÍM SE BUDEME ZABÝVAT PŘEVODEM NKA S ε-HRANAMI NA DETERMINISTICKÝ KA.

JSOU-LI V PŘECHODOVÉM GRAFU e-HRANY, MUSÍME PRO KAŽDÝ STAV VYTVOŘIT MNOŽINU STAVŮ, KTERÉ JSOU DOSAŽITELNÉ CESTAMI Z e-HRAN. PŘI KONSTRUKCI PŘECHODOVÉ TABULKY JIŽ ZNÁMÝM ALGORITMEM MUSÍME POČÍTAT S TÍM, ŽE MŮŽE-LI AUTOMAT PŘEJÍT DO STAVU X, MŮŽE PŘEJÍT I DO VŠECH JEHO e-NÁSLEDNÍKŮ.

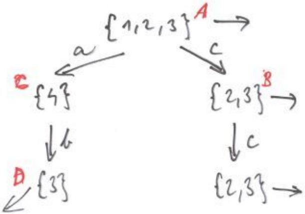
Př: $\Sigma = \{a, b, c\}$ $\quad R = ab + c^*$

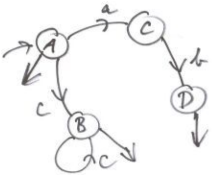




TABULKA e-NÁSLEDNÍKŮ

STAV	e-NÁSLEDNÍCI
→ 1	{1,2,3}
2	{2,3}
← 3	{3}
4	{4}





NÁSLEDUJÍCÍ TVRZENÍ JSOU EKVIVALENTNÍ:

JAZYK L JE TYPU 3.

JAZYK L LZE POPSAT REGULÁRNÍM VÝRAZEM.

EXISTUJE NEDETERMINISTICKÝ KA $\mathcal{A}$ TAKOVÝ, ŽE $L = L(\mathcal{A})$

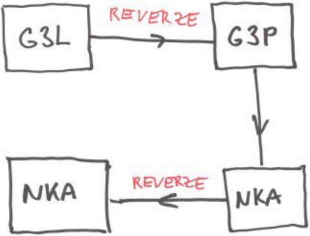
EXISTUJE DETERMINISTICKÝ KA $\mathcal{A}$ TAKOVÝ, ŽE $L = L(\mathcal{A})$.

RŮZNÉ REPREZENTACE JAZYKŮ TYPU 3

$$\begin{tikzcd}
\boxed{G3} \arrow[r] & \boxed{G3R} \arrow[r, leftrightarrow] & \boxed{NKA} \arrow[r] \arrow[d, leftrightarrow] & \boxed{KA} \\
 & & \boxed{RV} &
\end{tikzcd}$$

PROBRANÉ ALGORITMY PŘEVODU

JAK ZKONSTRUOVAT NKA K LEVÉ LINEÁRNÍ GRAMATICE ?



REVERZE GRAMATIKY:

REVERZE VŠECH LEVÝCH STRAN PŘEPISOVACÍCH PRAVIDEL.

REVERZE NKA:

ZMĚNA ORIENTACE VŠECH PŘECHODŮ VZÁJEMNÁ ZÁMĚNA VŠECH VSTUPNÍCH A VÝSTUPNÍCH STAVŮ.