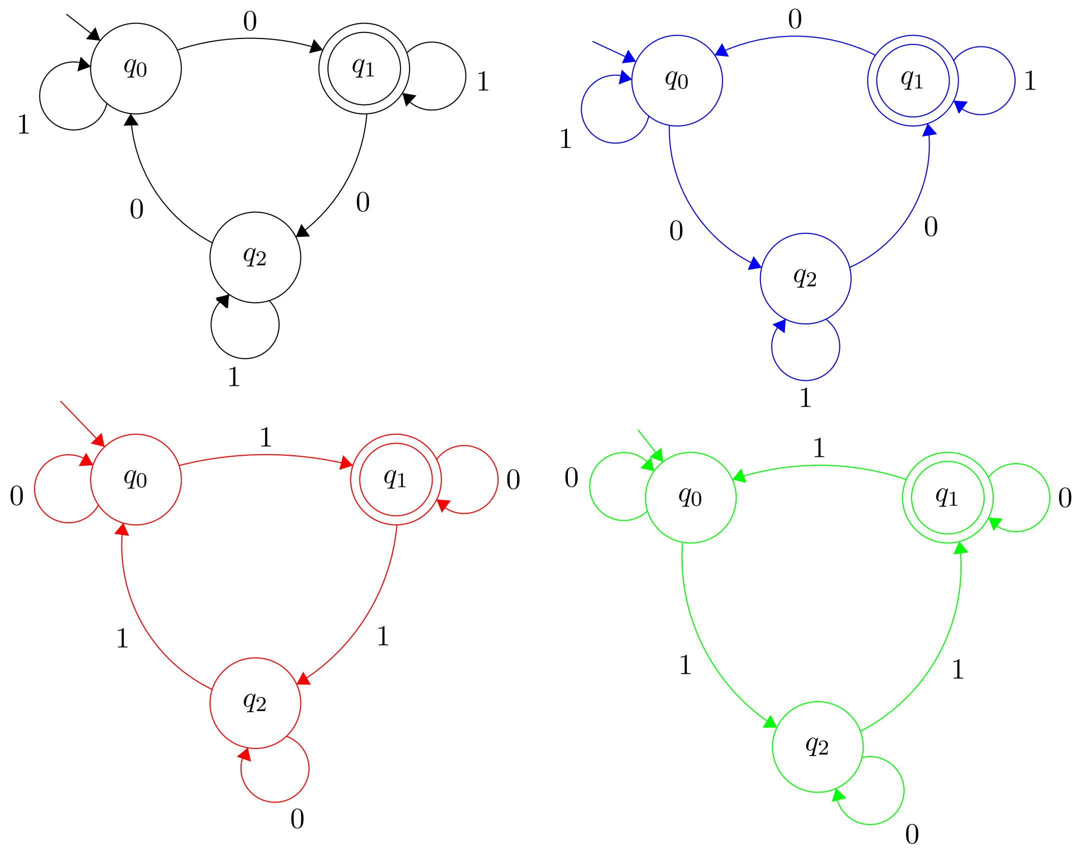
1. príklad

Na obrázku sa nachádzajú 4 automaty rozoznávajúce rôzne jazyky. Pre každý z nich zistite, aký jazyk rozoznáva a určite KL množiny všetkých stavov. Čo znázorňujú dolné indexy stavov v čiernom a červenom automate?



Riešenie príkladu 1

Všetky štyri automaty pracujú nad abecedou $\Sigma = \{0, 1\}$.

(a) **Čierny automat** rozoznáva jazyk $L_{black} = \{w \in \Sigma^* \mid |w|_0 \; mod \; 3 = 1\}$

- $KL[q_0] = \{\lambda;\; w \in \Sigma^* \mid |w|_0 \; mod \; 3 = 0\}$
- $KL[q_1] = \{w \in \Sigma^* \mid |w|_0 \; mod \; 3 = 1\}$
- $KL[q_2] = \{w \in \Sigma^* \mid |w|_0 \; mod \; 3 = 2\}$

(b) **Modrý automat** rozoznáva jazyk $L_{blue} = \{w \in \Sigma^* \mid |w|_0 \ mod\ 3 = 2\}$

- $KL[q_0] = \{\lambda;\ w \in \Sigma^* \mid |w|_0 \ mod\ 3 = 0\}$
- $KL[q_1] = \{w \in \Sigma^* \mid |w|_0 \ mod\ 3 = 2\}$
- $KL[q_2] = \{w \in \Sigma^* \mid |w|_0 \ mod\ 3 = 1\}$

(c) **Červený automat** rozoznáva jazyk $L_{red} = \{w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 1\}$

- $KL[q_0] = \{\lambda;\ w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 0\}$
- $KL[q_1] = \{w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 1\}$
- $KL[q_2] = \{w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 2\}$

(d) **Zelený automat** rozoznáva jazyk $L_{green} = \{w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 2\}$

- $KL[q_0] = \{\lambda;\ w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 0\}$
- $KL[q_1] = \{w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 2\}$
- $KL[q_2] = \{w \in \Sigma^* \mid |w|_1 \ mod\ 3 = 1\}$

Vidíme, že indexy stavov v červenom a čiernom automate značia zvyšky po delení. Samozrejme, aj modrý a zelený automat je správny, ale pre ľahšiu orientáciu a pre dokazovanie je vhodné používať jasné označovanie, ktoré vám priamo napovedá, čo sa v stave deje. Uvidíme to v ďalších príkladoch.

Zamyslite sa ešte nad tým, či je potrebné písať λ pri tých o KL množinách. Svoju odpoveď skúste zdôvodniť.

2. príklad

(a) Zostrojte automat, ktorý rozoznáva jazyk $L = \{w = ax \mid x \in \{a, b\}^* \wedge |w|_b \ mod\ 2 = 0\}$. Z akých dvoch jazykov je L zložený? Pre každý zostrojte automat, dokážte ich správnosť, a pomocou modulárnej konštrukcie (simulácie) vytvorte automat pre jazyk L.

(b) Vytvorte automat rozoznávajúci čísla deliteľné 6. Slová budú čísla:

(i) v dvojkovej sústave,

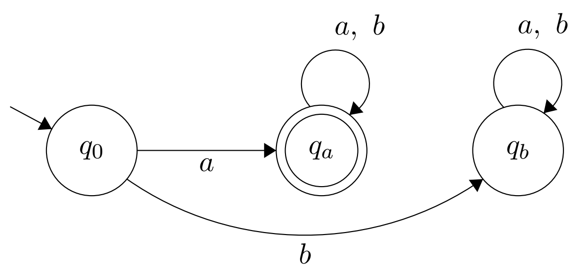
(ii)* v desiatkovej sústave.

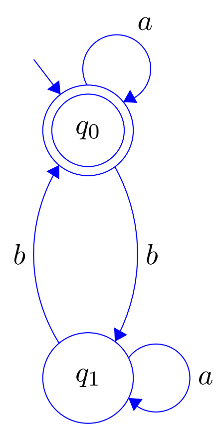
Kedy je číslo deliteľné 6? Vytvorte automaty pre oba prípady, overte ich správnosť a pomocou modulárnej konštrukcie vytvorte výsledný automat.

Riešenie príkladu 2a

Tento jazyk sa skladá z dvoch jazykov, označme ich L_1 a L_2. $L_1 = \{w = ax \mid x \in \{a,b\}^*\}$; $L_2 = \{w \in \{a,b\}^* \mid |w|_b \ mod\ 2 = 0\}$.

Najprv si pre každý z nich vytvoríme automat. Aby sa nám s automatmi ľahšie ďalej pracovalo, budeme sa snažiť označovať stavy podľa toho, čo sa v nich deje, a tiež sa vyvarovať veľa stavom s rovnakými menami.





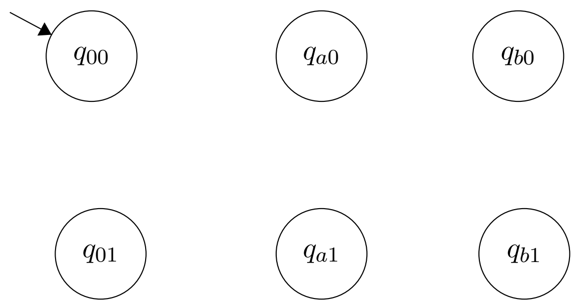
Vznikli dva automaty, pre každý si definujme KL triedy.

- **Čierny automat:** $L_1 = \{ax \mid x \in \{a,b\}^*\}$
  - $KL[q_0] = \{\lambda\}$
  - $KL[q_a] = \{ax \mid x \in \{a,b\}^*\}$ – čiže slová začínajúce znakom a
  - $KL[q_b] = \{bx \mid x \in \{a,b\}^*\}$ – čiže slová začínajúce znakom b
- **Modrý automat:** $L_2 = \{w \in \{a,b\}^* \mid |w|_b \ mod\ 2 = 0\}$
  - $KL[q_0] = \{w \in \{a,b\}^* \mid |w|_b \ mod\ 2 = 0\}$ – párny počet béčok
  - $KL[q_1] = \{w \in \{a,b\}^* \mid |w|_b \ mod\ 2 = 1\}$ – nepárny počet béčok

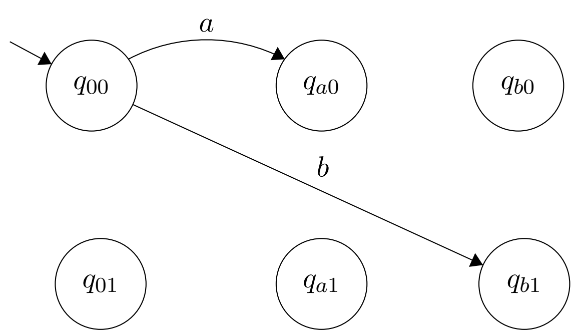
Dôkazy správnosti nechávam na vás ako opakovanie posledného cvičenia.

Poďme si spraviť modulárnu konštrukciu. Možno ste rozmýšľali, prečo má jeden automat všetky stavy v jednom riadku a druhý v jednom stĺpci. Je to preto, aby sa nám ľahko spravil kartézsky súčin stavov. Postup je nasledovný:

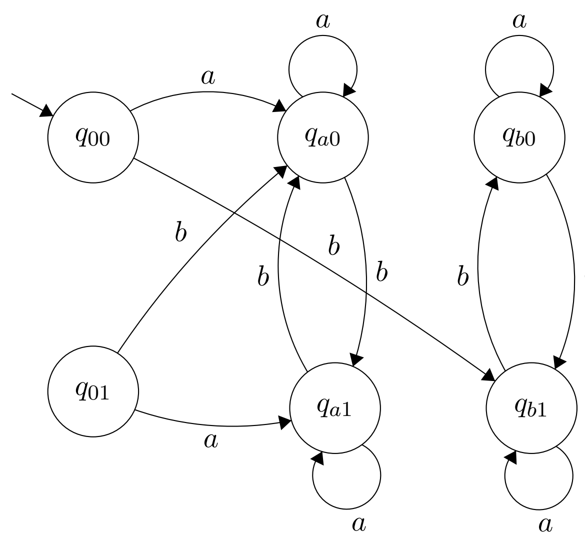
(i) Potrebujeme 3×2 stavov, každý bude mať dvojznakový index, kde prvý znak bude z čierneho automatu a druhý z modrého. Vieme, že vstupný stav bude ten, ktorý bol vstupný v oboch, v tomto prípade teda q_{00}.



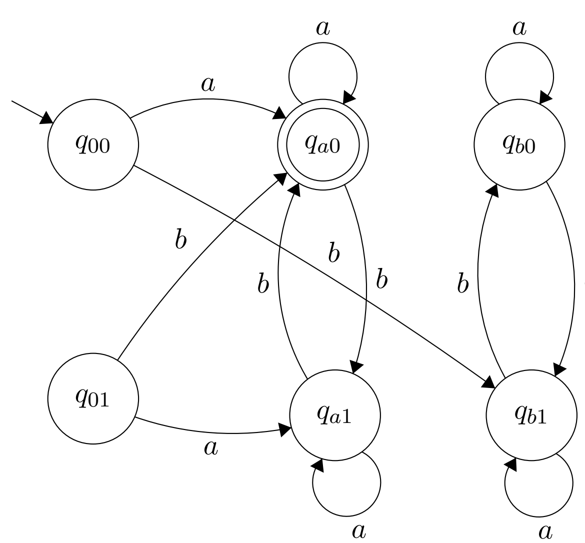
(ii) Teraz potrebujeme pridať prechody. Ako na to? Pozrime sa na prechody zo stavu q_{00}. V čiernom automate sa po a dostaneme zo stavu q_0 do stavu q_a a v modrom automate sa z q_0 dostaneme po a do q_0. Takže zo stavu q_{00} po a pôjde prechod do stavu q_{a0}. Podobne pre prechod po b – v čiernom automate sa zo stavu q_0 dostaneme do stavu q_b a v modrom automate sa z q_0 dostaneme do q_1. Z toho vyplýva, že zo stavu q_{00} po b pôjde prechod do stavu q_{b1}.



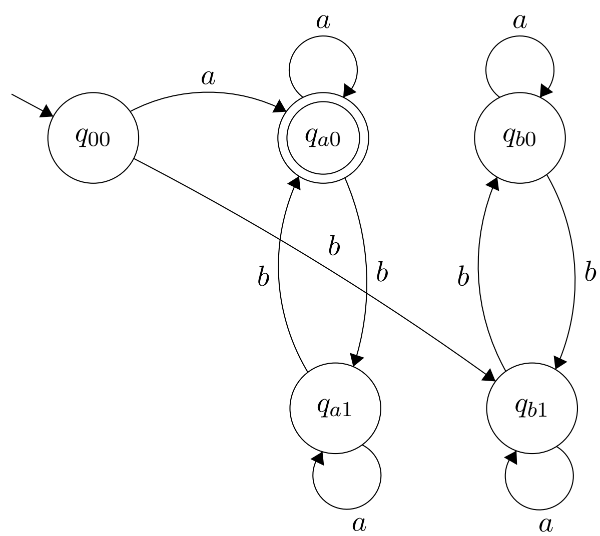
(iii) Všeobecnejšie: aby sme zistili prechod zo stavu q_{ij} po znaku e, tak najprv pozrieme, kam sa dostaneme zo stavu q_i po e v prvom automate (napr. do stavu q_k) a potom kam sa dostaneme zo stavu q_j po e v druhom automate (napr. do stavu q_l). Kombináciou výsledkov dostaneme výsledný stav (q_{kl}). Doplňme aj zvyšok prechodov v našom automate.



(iv) Keď už máme všetky prechody, ešte musíme určiť, ktoré stavy sú akceptačné. Keďže robíme prienik dvoch automatov (vo výslednom jazyku je medzi L_1 a L_2 „a zároveň”), akceptačné stavy budú tie, ktoré sú akceptačné aj v jednom aj v druhom automate. V našom prípade to je iba stav q_{a0}.



(v) Pozrite sa na automat. Sú v ňom všetky stavy dosiahnuteľné (teda sa do nich vieme počas výpočtu dostať)? Vidíme, že do stavu q_{01} sa dostať nedá. Ale prečo? Ak si spomenieme na $KL[q_0]$ z čierneho a $KL[q_1]$ modrého automatu, získame odpoveď. Neexistuje totiž žiadne slovo, ktoré je zároveň prázdne a zároveň má nepárny počet béčok. Tento stav aj s prislúchajúcimi prechodmi preto nemusíme kresliť.



(vi) Ešte si definujme KL triedy tohto automatu.

- $KL[q_{00}] = \{\lambda\}$
- $KL[q_{a0}] = \{w = ax \mid |w|_b \ mod\ 2 = 0\}$ – slová začínajúce a s párnym počtom b
- $KL[q_{b0}] = \{w = bx \mid |w|_b \ mod\ 2 = 0\}$ – slová začínajúce b s párnym počtom b
- $KL[q_{a1}] = \{w = ax \mid |w|_b \ mod\ 2 = 1\}$ – slová začínajúce a s nepárnym počtom b
- $KL[q_{b1}] = \{w = bx \mid |w|_b \ mod\ 2 = 1\}$ – slová začínajúce b s nepárnym počtom b

Pozrite sa na KL triedy pôvodného modrého a čierneho automatu vyššie. Ako súvisia s KL triedami tohto automatu? Vidíte tam prienik jazykov? Ako by sa zmenil automat, keby sme nerobili prienik, ale zjednotenie (t.j. v pôvodnom jazyku by bolo „alebo")? Zmenili by sa iba akceptačné stavy. Zamyslite sa, ktoré ďalšie stavy by boli v tomto prípade akceptačné.

Ak ste dokázali správnosť oboch automatov čierneho aj modrého, tak správnosť tohto už dokazovať nemusíte.

3. príklad

(a)* Zostrojte automat pre jazyk $L = \{axb \mid x \in \{a,b\}^*\}$.

(i) Dokážte jeho správnosť a určite KL triedy.

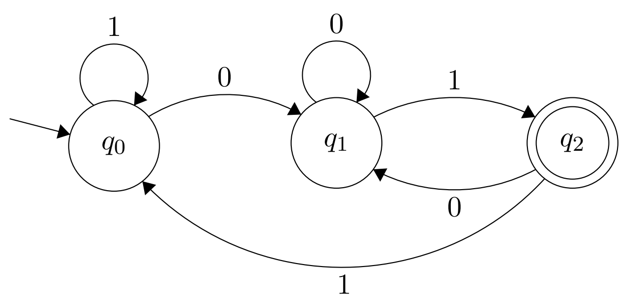
(ii) Aký minimálny počet stavov musí mať? Dokážte.

(b) Pozrite sa na automat A na obrázku a:

(i) zistite, aký jazyk rozoznáva,

(ii) určite KL triedy všetkých stavov,

(iii) dokážte, že automat rozoznávajúci jazyk na obrázku musí mať aspoň tri stavy.



Riešenie príkladu 3b

Poďme postupne po podúlohách.

(i) Ako zistiť, aký jazyk automat rozoznáva? Skúsme si vypísať niekoľko slov, ktoré skončia v akceptačnom stave.

Sú to napríklad slová: $01, 101, 1101, 0001, 1001, 0101....$ Čo majú tieto slová spoločné? Všetky z nich majú sufix 01. Ale budú ho mať všetky slová, ktoré skončia v q_2?

Pozrime sa, ako sa do q_2 vieme dostať. Jediný prechod, ktorý v q_2 končí, je prechod z q_1 po 1. Takže slová určite budú končiť 1. Do q_1 sa zas vieme dostať iba po 0 a to zo všetkých troch stavov. Takže áno, všetky slová končiace v q_2 musia obsahovať na konci 01. Najkratšie akceptované slovo je práve 01, výsledný jazyk môžeme zapísať ako:

$$L(A) = \{w01 \mid w \in \{0,1\}^*\}$$

(ii) Máme jazyk, takže máme aj $KL[q_2]$, poďme zistiť zvyšné dve. Je vhodné zachovávať rovnaký zápis tried, preto sa pri stavoch q_0 a q_1 zamyslime, aký sufix budú mať slová, ktoré v nich skončia.

Jednoduchší je asi stav q_1, lebo doňho vedie iba jeden prechod a to z q_0 po 0. Takže v q_1 skončia slová so sufixom 0. Je táto množina disjunktná s $KL[q_2]$ čiže $L(A)$? Áno. Zapíšme ju.

$$KL[q_1] = \{w0 \mid w \in \{0,1\}^*\}$$

No a čo so stavom q_0? Vieme, že tam skončí λ, slová z $\{1\}^*$, no ale aj také, čo prejdú všetkými stavmi a z q_2 sa vrátia po 1. Najjednoduchšie by bolo povedať, že sú to slová so sufixom 1, ale bola by takáto množina disjunktná s $KL[q_2]$? Alebo inak – existuje slovo, ktoré má aj sufix 01, aj 1? Samozrejme, napríklad najkratšie také slovo 01. Takže zjavne tieto množiny disjunktné nie sú.

Čím budú končiť slová, ktoré sa do q_0 dostanú z q_2? 1 a pred ňou môže byť iba 1, keďže to je jediný prechod do q_2, takže tieto slová budú mať sufix 11. Zapíšme to všetko.

$$KL[q_0] = \{\lambda\} \cup \{1^n \mid n \in N\} \cup \{w11 \mid w \in \{0,1\}^*\}$$

Vieme tento zápis zjednodušiť? Ktoré slová z prvých dvoch jazykov nepokrýva posledný jazyk? Iba $\lambda, 1$. Takže to upravíme a napíšeme všetky KL triedy:

$$KL[q_0] = \{\lambda, 1; w11 \mid w \in \{0,1\}^*\}$$
$$KL[q_1] = \{w0 \mid w \in \{0,1\}^*\}$$
$$KL[q_2] = \{w01 \mid w \in \{0,1\}^*\}$$

(iii) Máme dokázať, že automat rozoznávajúci tento jazyk musí mať aspoň tri stavy. Dokazujeme sporom, t. j. predpokladáme, že automat môže mať aj menej ako tri stavy. (Čo to znamená? Že vieme nejaké dva stavy zlúčiť. Poďme ukázať, že to nejde.)

Z každej KL triedy si zvoľme ľubovoľné slovo, napríklad $011, 00, 101$. Pre každú dvojicu ideme zistiť, či môžu skončiť v rovnakom stave.

(i) Nech $\hat{\delta}(q_0, 011) = \hat{\delta}(q_0, 00)$. Potom, z Lémy 3.12 musí platiť, že keď k nim pridáme hocijaké slovo, oba skončia v rovnakom stave, a teda buď obe slová budú z jazyka alebo obe slová nebudú z jazyka. Čiže $\hat{\delta}(q_0, 011x) = \hat{\delta}(q_0, 00x)$, $011x \in L(A) \iff 00x \in L(A)$.

Aby sme ukázali, že tieto slová nemôžu skončiť v jednom stave, hľadáme x, ktoré túto rovnosť poruší. Existuje také?

Skúsme napríklad iba jednoduché slovo 1. $\hat{\delta}(q_0, 0111) = \hat{\delta}(q_0, 001)$. Vidíme, že $0111 \notin L(A)$ ale $001 \in L(A)$. Z toho vyplýva, že slová 011 a 00 nemôžu skončiť v jednom stave a preto nevieme zlúčiť stavy q_0 a q_1.

(ii) Nech $\hat{\delta}(q_0, 011) = \hat{\delta}(q_0, 101)$. Tu vidíme, že už toto tvrdenie nemôže platiť, keďže $011 \notin L(A)$ ale $101 \in L(A)$. Tieto slová nemôžu skončiť v rovnakom stave (t.j. nevieme zlúčiť stavy q_0 a q_2.)

(iii) Posledná kombinácia: nech $\hat{\delta}(q_0, 00) = \hat{\delta}(q_0, 101)$. Opäť neplatí priamo toto tvrdenie, pretože $00 \notin L(A)$ ale $101 \in L(A)$. Tieto slová tak nemôžu skončiť v rovnakom stave (t.j. nevieme zlúčiť stavy q_1 a q_2.)

Dospeli sme k sporu s tvrdením, že automat môže mať menej ako tri stavy, takže platí pôvodné tvrdenie. Automat rozoznávajúci jazyk $L(A) = \{w01 \mid w \in \{0,1\}^*\}$ musí mať aspoň 3 stavy.

4. príklad

Dokážte, že nasledujúce jazyky nie sú regulárne (s využitím Lémy 3.12 z prednášky).

(a) $L_1 = \{0^n 1^n \mid n \in \mathbb{N}\}$

(b)* $L_2 = \{w\#w \mid w \in \{a,b\}^*\}$; $\Sigma = \{a, b, \#\}$

(c)* $L_3 = \{ww^R \mid w \in \{a,b\}^*\}$

Riešenie príkladu 4a

Čo znamená, že je jazyk regulárny? Že vieme zostrojiť konečný stavový automat, ktorý bude tento jazyk rozpoznávať. Ak je jazyk neregulárny, takýto automat zostrojiť nevieme. Zamyslime sa, či budeme vedieť zostrojiť automat pre jazyk L_1. Čo by sme si museli pamätať?

Najprv by sme potrebovali mať stavy, v ktorých si pamätáme, koľko núl sme prečítali. Potom, z každého takéhoto stavu by sme išli do stavov, v ktorých by sme si počítali počet jednotiek. Je ale možné takýto automat zostrojiť pre $n \in \mathbb{N}$? Vieme povedať, aký by bol najväčší počet núl a jednotiek? Nekonečno. Náš automat má však iba konečný počet stavov, čiže tento jazyk nevie rozoznávať, a preto je neregulárny.

Dokážme to sporom, čiže dokazujeme opačné tvrdenie a to, že jazyk L_1 je regulárny. Potom, pre takýto jazyk vieme zostrojiť deterministický automat A, ktorý má konečný počet stavov, čiže $|Q| = m$. Pre každý deterministický automat platí léma 3.12, ktorá hovorí, že ak dve slová skončia v rovnakom stave a prireťazíme k nim obom rovnaké slovo, tak obe tieto nové slová buď budú z jazyka $L(A)$, alebo nebudú z $L(A)$.

Kedy dve slová určite skončia v rovnakom stave? Ak ich je viac, ako stavov. Vyberme teda takú množinu slov, ktorá bude aspoň o jedna väčšia, ako počet stavov m. Nech sú to slová tvaru 0^k, kde $1 \leq k \leq m+1$. Vyberme z nich dve rôzne, ktoré skončia v rovnakom stave. Nevieme povedať, ktoré presne to budú, preto budeme používať premenné. Keďže nebudú rovnaké, jedno bude dlhšie ako druhé.

Nech $\hat{\delta}(q_0, 0^i) = \hat{\delta}(q_0, 0^j)$, kde $1 \leq i < j \leq m+1$.

Potom z Lémy 3.12 vyplýva, že $0^i z \in L(A) \Leftrightarrow 0^j z \in L(A)$, pre každé $z \in \Sigma^*$ (pretože ak sme pri čítaní 0^i a 0^j skončili v rovnakom stave, keď z tohto stavu začneme čítať z, vieme skončiť iba v jednom stave, čiže $\hat{\delta}(q_0, 0^i z) = \hat{\delta}(q_0, 0^j z)$. A ak obe skončia v rovnakom stave a jedno by bolo z jazyka a druhé nie, vedeli by sme o tomto stave povedať, či je akceptačný?).

Teraz si musíme zvoliť z tak, aby jedno slovo bolo z jazyka L_1 a druhé nie, tak získame spor s tvrdením, že jazyk je regulárny. Ako vyzerajú slová z jazyka? Majú niekoľko núl a po nich rovnaký počet jednotiek. Zvoľme preto $z = 1^i$. Potom:

- $0^i 1^i \in L(A)$
- $0^j 1^i \notin L(A)$, keďže $i < j$, čiže v tomto slove bude viac núl ako jednotiek.

Dostali sme sa do sporu s tvrdením, že jazyk L_1 je regulárny, platí preto pôvodné tvrdenie – jazyk L_1 je neregulárny.

Zamyslite sa, či aj jazyk $L = \{0^n 1^n \mid 0 < n < 2021\}$ je neregulárny, alebo preň vieme zostrojiť automat, ktorý ho bude rozpoznávať.

Riešenie príkladu 4b

Dokazujeme sporom. T.j. dokazujeme tvrdenie, že jazyk $L_2 = \{w\#w \mid w \in \{a,b\}^*\};\ \Sigma = \{a,b,\#\}$ je regulárny. Potom musí existovať automat $A = \{Q, \Sigma, \delta, q_0, F\}$, pre ktorý platí že $L(A) = L_2$. Nech počet stav tohto (konečného) automatu je n.

Keď vezmeme slová $a, a^2, a^3, ..., a^{n+1}$, tak potom[1] musia aspoň dve slová skončiť v rovnakom stave, čiže $\hat{\delta}(q_0, x) = \hat{\delta}(q_0, y)$.

Nech $x = a^i$ a $y = a^j$, pričom $1 \leq i < j \leq n+1$. (Sú to teda nejaké dve rôzne slová z postupnosti vyššie[2].) Čiže $\hat{\delta}(q_0, a^i) = \hat{\delta}(q_0, a^j)$. A z Lémy 3.12 vieme, že potom platí $a^j z \in L(A) \Leftrightarrow a^i z \in L(A)$ (pretože $\hat{\delta}(q_0, a^j z) = \hat{\delta}(q_0, a^i z)$).

Zvoľme si $z = \#a^i$. Takže $\hat{\delta}(q_0, a^j\#a^i) = \hat{\delta}(q_0, a^i\#a^i)$.

- $a^j\#a^i \notin L(A)$, pretože $i < j$
- $a^i\#a^i \in L(A)$

Dostali sme sa do sporu. Preto platí pôvodné tvrdenie a jazyk L_2 nie je regulárny.

Mohli by ste sa pýtať, prečo sme zvolili a^i, a^j a nikde v slove sa nenachádza b. Je na nás, akú množinu slov si zvolíme, nemusíme použiť všetky znaky abecedy, ale na konci potrebujeme dostať spor – čiže ak k našim slovám prireťazíme nejaké iné slovo, jedno bude z pôvodného jazyka a druhé nie. Čo okrem $\#a^i$ by sme mohli k slovám ešte prireťaziť, aby sme sa dostali do sporu? Napríklad $b^3\#a^ib^3$, $ba\#a^iba$... Čiže to, čo dáme pred mriežku, zopakujeme aj za ňou.

Čo by sa stalo, ak by sme prireťazili $\#a^j$? Dostali by sme spor? Áno, keďže slovo $a^j\#a^j \in L(A)$, ale $a^i\#a^j \notin L(A)$.

Poznámky

- Dôkaz minimálneho počtu stavov pre automat:
  - Z každej KL triedy (pre každý stav) si zvolíme jedno slovo, ktoré do nej patrí.

[1] z Pigeonhole principle

[2] Častou chybou je, že vyberiete nejaké konkrétne slová. To spraviť nemôžeme, lebo nevieme, ktoré z nich skončia v rovnakom stave. Preto ich musíme zapísať takto všeobecnejšie a dávať pritom pozor, aby sme zaručili, že tie slová budú rôzne a z množiny väčšej ako je počet stavov automatu.

  - Dokazujeme sporom: *Automat môže mať aj menej ako n stavov.*
  - Využijeme Lému 3.12. Pre každú dvojicu slov zisťujeme, či môžu skončiť v rovnakom stave. Môžu nastať dva prípady:
    i) jedno slovo do jazyka patrí a druhé nie, takže priamo sme dokázali, že nemôžu skončiť v rovnakom stave,
    ii) alebo po „prilepení" vhodného sufixu bude jedno slovo patriť do jazyka, ale druhé nie.
  - Keď nám pre všetky dvojice (tj. pre $\binom{n}{2}$ kombinácii) vyjde, že slová nemôžu skončiť v rovnakom stave, dospeli sme k sporu s tvrdením a teda platí pôvodné tvrdenie *Automat musí mať aspoň n stavov.*
  - Ak náhodou vyjde, že nejaké dve slová by mohli skončiť v rovnakom stave, skúste automat prerobiť - možno môže byť aj menší ako ste si pôvodne mysleli :).

- Modulárna konštrukcia
  - Najprv treba zostrojiť oba automaty, z ktorých budeme skladať výsledný.
  - Potom dokázať správnosť oboch automatov – t.j. že rozoznávajú jazyk, ktorý majú.
  - Keď sme to dokázali, spravíme modulárnu konštrukciu (simuláciu). Budeme potrebovať $n \times m$ stavov, ak n je počet stavov jedného automatu a m je počet stavov druhého. Odporúčam nové stavy indexovať kombináciou indexov z pôvodných automatov.
  - Vstupným stavom bude taký stav, ktorý bol v oboch automatoch vstupný.
  - Kam ide prechod zo stavu q_{ij} po znaku e zistíme tak, že sa najprv pozrieme, kam sa dostaneme zo stavu q_i po e v prvom automate (napr. do stavu q_k) a potom kam sa dostaneme zo stavu q_j po e v druhom automate (napr. do stavu q_l). Kombináciou výsledkov dostaneme výsledný stav (q_{kl}).
  - Akceptačné stavy budú:
    (a) Ak robíme *prienik* automatov, tak všetky ktoré sú akceptačné pre *oba* automaty.
    (b) Ak robíme *zjednotenie* automatov, tak všetky, ktoré sú akceptačné *aspoň pre jeden* automat.

    (c) Ak robíme *rozdiel* automatov, tak všetky, pre ktoré je *prvý akceptačný a druhý nie.*

  - Výsledný automat už dokazovať nemusíme.

- Dôkaz neregulárnosti pomocou Lémy 3.12
  - Dokazujeme sporom, takže predpokladáme, že jazyk L je regulárny a teda existuje DKA A, ktorý ho rozoznáva $L = L(A)$.
  - Ak máme konečný automat s n stavmi, tak ak budeme mať čo i len o jedno slovo viac, ako je stavov $(n + 1)$, musia výpočty na aspoň dvoch slovách skončiť v rovnakom stave (Pigeonhole principle).
  - Zvolíme si nejakú množinu slov, ktorých je dostatočne veľa (aspoň $n + 1$ pre ľubovoľné n), takže napr. 0^n, a^{2k}... Treba vysvetliť/ukázať, prečo je ich určite viac ako stavov automatu. Z nich vyberiete dve slová. Pozor, nie ľubovoľné! Treba si uvedomiť, že vy neviete, ktoré konkrétne dve slová skončia v rovnakom stave, ale viete, že ak ste vytvorili aspoň $n + 1$ slov, tak to mohlo byť ľubovolné i-te a j-te. Preto *nevhodné slová* sú také, kde používate konštanty a taktiež také, kde jedno závisí od druhého napr. $1^i, 1^{2i}$ alebo a^5, a^{5+k}.
  - Z Lémy 3.12 vieme, že ak výpočet na slovách u a v skončí v rovnakom stave q, tak aj výpočty na ux a vx skončia v jednom stave p, čiže slovo ux patrí do jazyka L práve vtedy keď aj slovo vx.
  - Slovo x je preto vhodné vybrať tak, aby jedno slovo z dvojice ux, vx patrilo do jazyka a druhé nie. Stav p by tak mal jedno slovo akceptovať a druhé nie, to DKA robiť nevie. Takže buď ich obe akceptuje alebo nie, čiže nerozoznáva daný jazyk, čím sme sa dostali do sporu s tvrdením, že jazyk je regulárny.
  - Odporúčam prečítať si komentáre k úlohe 5.2 z roku 2017/18