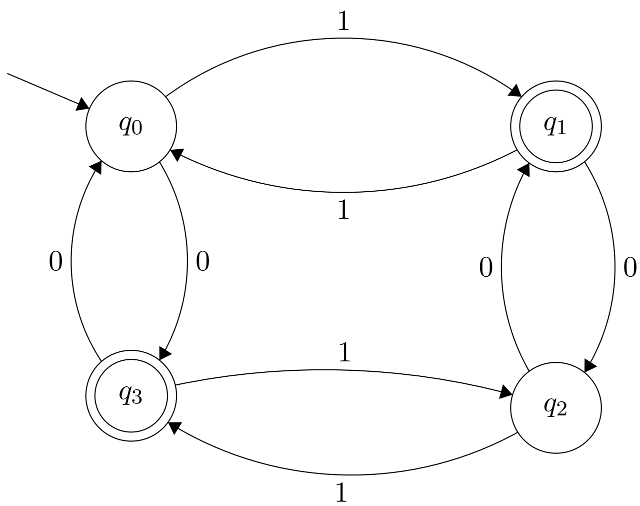
1. príklad

Aký jazyk rozpoznáva automat na obrázku? Dokážte to.



Idea riešenia 1. príkladu

Skúsme zistiť, aký jazyk rozoznáva daný automat. V stave q_0 skončí určite prázdne slovo. Koľko jednotiek bude obsahovať slovo, ktoré skončí v q_0? Vieme sa tam dostať cez stav q_1, pričom na ceste $q_0 \longrightarrow q_1 \longrightarrow q_0$ bude vždy párny počet jednotiek, hocikoľkokrát túto cestu zopakujeme. Môžeme prejsť až do q_2, ale všimnime si, že či sa do q_0 vrátime zo stavu q_3 alebo q_1, vždy prejdeme po párnom počte jednotiek (jednou vychádzame a druhou vchádzame). Rovnako to je pre q_0 s nulami – buď pôjdeme do q_3 a späť, čo vždy dá párny počet núl, alebo pôjdeme cez stav q_2 a opäť budeme potrebovať párny počet núl. V stave q_0 by tak mali končiť slová, ktoré majú párny počet núl a párny počet jednotiek. To vieme zapísať ako $KL[q_0] = \{w \mid |w|_0 \; mod \; 2 = 0 \wedge |w|_1 \; mod \; 2 = 0\}$

Ako sa vieme dostať do q_1? Buď priamo z q_0 a na ceste prečítame jednu jednotku, alebo pôjdeme cez q_3 a na ceste prečítame jednu jednotku a dve nuly. Hocijaký iný okruh, ktorý z q_1 spravíme, bude obsahovať párny počet núl aj jednotiek. To spočítané s tými, ktoré sme spravili pri prvom prechode do q_1 nám dáva nepárny počet jednotiek a párny počet núl. Čiže: $KL[q_1] = \{w \mid |w|_0 \; mod \; 2 = 0 \wedge |w|_1 \; mod \; 2 = 1\}$.

Do q_3 sa vieme dostať priamo z q_0 po jednej nule, alebo cez zvyšné stavy po dvoch jednotkách a jednej nule. Tiež tu platí to, čo v predchádzajúcich stavoch, ak z q_3 pôjdeme do iného stavu a späť, prečítame párny počet núl a/alebo párny počet jednotiek. Z toho vyplýva, že v tomto stave skončia slová, ktoré majú nepárny počet núl a párny počet jednotiek. $KL[q_3] = \{w \mid |w|_0 \ mod\ 2 = 1 \wedge |w|_1 \ mod\ 2 = 0\}$.

A čo stav q_2? Keď sa pozrieme na KL triedy zvyšných stavov, vidíme že nám chýbajú ešte slová, ktoré majú nepárny počet núl a nepárny počet jednotiek. Skontrolujme. Do q_2 sa vieme dostať cez q_1 prečítaním jednej jednotky a jednej nuly, alebo cez q_3 tiež prečítaním jednej nuly a jednej jednotky. Takže $KL[q_2] = \{w \mid |w|_0 \ mod\ 2 = 1 \wedge |w|_1 \ mod\ 2 = 1\}$.

Z toho vyplýva, že $L(M) = \{w \mid |w|_0 \ mod\ 2 = 0 \wedge |w|_1 \ mod\ 2 = 1\}$.

Dokážte, že to platí.

2. príklad

Zostrojte DKA, ktorý rozoznáva jazyk L a dokážte jeho správnosť/dokážte, že ste zostrojili automat s minimálnym počtom stavov.

(a) $L_1 = \{ax \mid x \in \{a,b\}^*\}$

(b) $L_2 = \{1, 00, 01\}$; $L_2 \subseteq \{0,1\}^*$

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

(d) $L_4 = \{w \mid |w|_b \ mod\ 2 = 1; w \in \{a,b,c\}^*\}$

3. príklad

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

(i) Dokážte jeho správnosť a určite KL triedy.

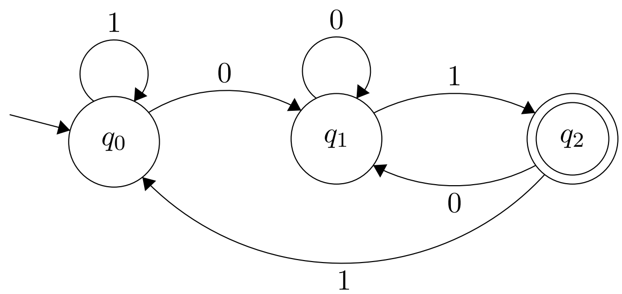
(ii) Aký minimálny počet stavov musí mať? Dokážte.

(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.



4. príklad

Vieme pre nasledujúce jazyky zostrojiť automat?

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

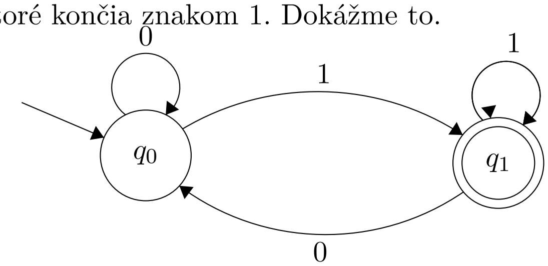
(b) $L_2 = \{w\#w \mid w \in \{a,b\}^*\};\ \Sigma = \{a,b,\#\}$

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

Skúste zostrojiť automat pre tieto jazyky, ako bude platiť, že $|w| \leq 3$ resp. $n \leq 3$.

5. príklad*

Na obrázku je automat, ktorý rozpoznáva slová nad abecedou $\{0,1\}$, ktoré končia znakom 1. Dokážme to.



Riešenie 5. príkladu

V stave q_0 skončia slová, ktoré končia znakom 0 a v stave q_1 skončia slová, ktoré končia znakom 1. Zapíšme to formálne:

- $KL[q_0] = \{\lambda, w0 \mid w \in \Sigma^*\}$
- $KL[q_1] = \{w1 \mid w \in \Sigma^*\}$

Toto bude našou hypotézou, ktorú budeme overovať pomocou matematickej indukcie (MI). [1]

Báza indukcie Spravíme výpočet na automate pre slová dĺžky 0 a vyššie, pokiaľ sa nedostaneme do všetkých stavov.

- *slová dĺžky 0:* $\hat{\delta}(q_0, \lambda) = q_0 \implies \lambda \in KL[q_0]$.
- *slová dĺžky 1:*
  - $\hat{\delta}(q_0, 0) = q_0 \implies 0 \in KL[q_0]$.
  - $\hat{\delta}(q_0, 1) = q_1 \implies 1 \in KL[q_1]$.

Obe slová 0 a 1 vieme napísať ako $\lambda 0$ a $\lambda 1$, čiže je zrejmé, že patria do KL tried stavov, v ktorých skončili. Dostali sme sa do všetkých stavov, báza indukcie je preto kompletná.

Indukčný predpoklad: Predpokladáme, že hypotéza platí pre všetky slová dĺžky $k, k \leq 1$. Alebo zapísané inak: slová $w \in \Sigma^*$, kde $|w| = k, k \leq 1$ (keďže platnosť hypotézy pre slová týchto dĺžok sme dokázali v báze).

Indukčný krok: Ak dokážeme, že naša hypotéza platí pre všetky slová dĺžky $k + 1$, kde $k \geq 1$, dokážeme všeobecnú platnosť našej hypotézy. Slová tejto dĺžky vieme zapísať ako $z = wa$, kde $w \in \Sigma^*, |w| = k$ a $a \in \Sigma$. O w môžu platiť dve veci - buď bude končiť 0 alebo bude končiť 1. Musíme ich obe rozobrať.

- w končí 0, takže z IP vieme, že patrí do $KL[q_0]$. Teraz nám opäť vzniknú dva prípady - $a = 0$ alebo $a = 1$.
  - $\hat{\delta}(q_0, w0) = \delta(\hat{\delta}(q_0, w), 0) = \delta(q_0, 0) = q_0$. Slovo $w0$ končí nulou, takže spĺňa našu hypotézu, že $z0 \in KL[q_0]$.
  - $\hat{\delta}(q_0, w1) = \delta(\hat{\delta}(q_0, w), 1) = \delta(q_0, 1) = q_1$. Slovo $w1$ končí jednotkou, takže spĺňa našu hypotézu, že $w1 \in KL[q_1]$.

[1] Ešte predtým sa ale musíme zamyslieť, či (1) KL triedy sú vzájomne disjunktné (zjavne slovo nevie končiť aj 0 aj 1, takže sú), či (2) každé slovo nad abecedou patrí do práve jednej KL triedy (máme tam priamo spomenuté aj prázdne slovo, zvyšné slová nad abecedou $\Sigma = \{0, 1\}$ vieme jednoznačne zaradiť do jednej KL triedy), či (3) zjednotenie všetkých KL tried bude obsahovať všetky slová z Σ^* (zjavne áno), a či (4) zjednotenie KL tried akceptačných stavov dá jazyk, ktorý má automat rozoznať (máme len jeden akceptačný stav, takže áno).

- w končí 1, takže z IP vieme, že patrí do $KL[q_1]$. Teraz nám opäť vzniknú dva prípady - $a = 0$ alebo $a = 1$.
  - $\hat{\delta}(q_0, w0) = \delta(\hat{\delta}(q_0, w), 0) = \delta(q_1, 0) = q_0$. Slovo $w0$ končí nulou, takže spĺňa našu hypotézu, že $z0 \in KL[q_0]$.
  - $\hat{\delta}(q_0, w1) = \delta(\hat{\delta}(q_0, w), 1) = \delta(q_1, 1) = q_1$. Slovo $w1$ končí jednotkou, takže spĺňa našu hypotézu, že $w1 \in KL[q_1]$.

Dokázali sme platnosť našej hypotézy. Čiže sme dokázali platnosť KL tried, a teda aj to, že jazyk rozpoznáva slová končiace 1 (tj. slová, ktoré skončia v stave q_1, čomu zodpovedá $KL[q_1]$).

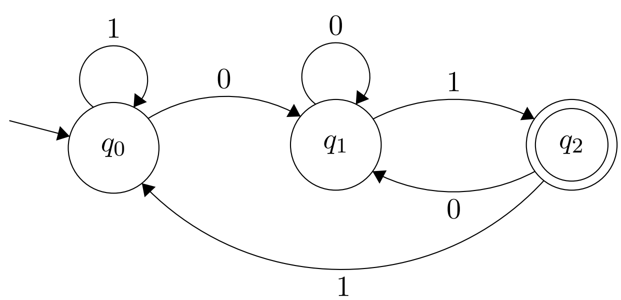
6. príklad*

Pozrite sa na automat A na obrázku a:

(a) zistite, aký jazyk rozoznáva,

(b) určite KL triedy všetkých stavov,

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



Riešenie príkladu 6

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.

Poznámky

KL triedy

- Musíme zadefinovať KL pre každý stav automatu.
- Zjednotenie KL tried pre všetky stavy sa musí rovnať Σ^*.

- Zjednotenie KL tried pre akceptačné stavy sa musí rovnať jazyku, ktorý automat rozoznáva.
- Teda - každé slovo nad abecedou musí patriť do nejakej KL triedy, a každé slovo z jazyka, ktorý automat rozoznáva, musí patriť do KL triedy jedného z akceptačných stavov.
- Prienik akýchkoľvek dvoch KL tried je prázdna množina - t.j. KL triedy sú disjunktné množiny.

MI pre automat

- Najprv si zadefinujeme KL triedy pre všetky stavy, to bude naša hypotéza (overte si všetko, čo musí platiť o KL triedach).
- V báze postupne na slovách dĺžky 0, 1... zisťujeme, kde jednotlivé slová v automate skončia, a či naozaj patria do KL triedy daného stavu. Dokazujeme po takú dĺžku slov, kým v báze nepokryjeme všetky stavy.
- Indukčný predpoklad je, že náš automat správne rozoznáva slová dĺžky $k \leq x$, kde x je (konkrétna) dĺžka najdlhšieho slova z bázy. $w, |w| = k, k \leq x$.
- Indukčný krok - dokazujeme pre slová dĺžky $k + 1$. $z = wa, |z| = k + 1, a \in \Sigma$.
- Dokazujeme pre všetky KL triedy, do ktorých môže patriť w a pre každú triedu ešte pre každý symbol z abecedy, ktoré k nemu vieme "prilepiť", aby sme mali slovo z.
- Treba si uvedomiť, že slovo w nielen končí v stave q_i, ale má aj vlastnosti $KL[q_i]$. Tie využijeme v IK a pre každú možnosť ukážeme, že ak k slovu w s nejakými vlastnosťami pridáme symbol a, slovo z skončí v stave q_j a aj vlastnosťami bude patriť do triedy $KL[q_j]$.
- Častou chybou je, že iba píšete, čo by malo pre dané slovo platiť, ale nekontrolujete si, či to aj naozaj platí. Potom si nenájdete chyby, či neuvedomíte, že ste na nejakú možnosť pri definovaní KL tried zabudli. Takže si to vždy overte, pri zložitejších triedach môžete začať na konkrétnych príkladoch, snažte sa ale dostať k všeobecnému opisu.

Dôkaz minimálneho počtu stavov pre automat:

- Z každej KL triedy (pre každý stav) si zvolíme jedno slovo, ktoré do nej patrí.

- 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.