Vzorové riešenia úlohy 5.1

Vzorové riešenie Máme daný jazyk:

$$L = \{w \in \{0,1\}^*, |w|_1 \mod 2 = 0, w = 10y00z, y, z \in \{0,1\}^*\}$$

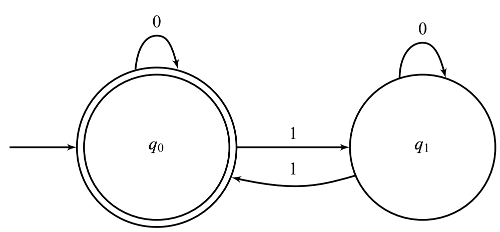
Pripomeňme si zadanie: máme napísať L ako prienik dvoch jazykov. Potom pre oba jazyky máme navrhnúť automat. Napokon máme všeobecnou metódou simulácie skonštruovať automat akceptujúci L. Zdôraznime niektoré detaily tohto zadania: oba čiastkové automaty **musíme** dokázať, výsledný automat **nemusíme, ak použijeme presne konštrukciu simulácie**, ktorej všeobecný dôkaz sa nachádza v knižke.

Ľahko vidno, že L sa dá rozpísať na nasledovné jazyky,

$$L_1 = \{w \in \{0,1\}, |w|_1 \mod 2 = 0\}$$
$$L_2 = \{w \in \{0,1\}, w = 10y00z, y, z \in \{0,1\}^*\}$$

pričom platí $L = L_1 \cap L_2$.

Skonštruujeme najprv M_1 taký, že $L(M_1) = L_1$. $M_1 = (K_1, \{0,1\}, \delta_1, q_0, F_1)$, kde $K_1 = \{q_0, q_1\}, F_1 = \{q_0\}$ a δ_1 je vidno z obrázka:



Poďme teraz dokázať správnosť M_1. Vyslovme najprv hypotézy ohľadom tried Kl:

$$Kl[q_0] = \{w \in \{0,1\}^*, |w|_1 \mod 2 = 0\}$$
$$Kl[q_1] = \{w \in \{0,1\}^*, |w|_1 \mod 2 = 1\}$$

Dokážeme správnosť týchto tried pre každé vstupné slovo matematickou indukciou od dĺžky vstupného slova w.

Báza indukcie: Ak $w = \lambda$, výpočet skonči v stave q_0. Ak $|w| = 1$, tak ak $w = 0$, výpočet skončí v q_0, ináč v q_1, čo zodpovedá našej hypotéze.

Indukčný krok: Ak $|w| = i > 1$, potom $w = ua$, kde $u \in \{0,1\}^{i-1}, a \in \{0,1\}$. Výpočet na w sa teda skladá z výpočtu na u a poslednom kroku na a:

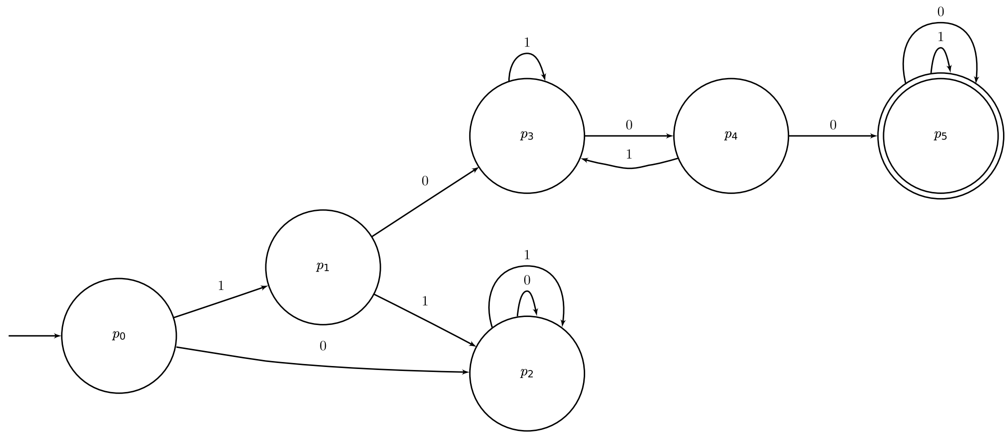
$(q_0, ua) \vdash^{i-1} (\hat{\delta_1}(q_0, u), a) \vdash (\delta_1(\hat{\delta_1}(q_0, u), a), \lambda)$.

Keďže $|u| < i$, potom sa na výpočet na u vzťahuje indukčný predpoklad a teda platí, že $\hat{\delta_1}(q_0, u) = q_j$, kde j je parita počtu jednotiek v u. Uvažujme 4 prípady:

- $\hat{\delta_1}(q_0, u) = q_0$, $a = 0$, potom $|w|_1 \mod 2 = 0$ a $\hat{\delta_1}(q_0, w) = q_0$.
- $\hat{\delta_1}(q_0, u) = q_1$, $a = 0$, potom $|w|_1 \mod 2 = 1$ a $\hat{\delta_1}(q_0, w) = q_1$.
- $\hat{\delta_1}(q_0, u) = q_0$, $a = 1$, potom $|w|_1 \mod 2 = 1$ a $\hat{\delta_1}(q_0, w) = q_1$.
- $\hat{\delta_1}(q_0, u) = q_1$, $a = 1$, potom $|w|_1 \mod 2 = 0$ a $\hat{\delta_1}(q_0, w) = q_0$.

Dokázali sme platnosť tried aj pre slová s dĺžkou i, čím sme dokončili dôkaz správnosti tried $Kl[q_0]$ a $Kl[q_1]$. Keďže $F_1 = \{q_0\}$, potom $L(M_1) = Kl[q_0] = \{w \in \{0,1\}^*, |w|_1 \mod 2 = 0\} = L_1$.

Navrhnime teraz automat M_2 taký, že $L(M_2) = L_2$. $M_2 = (K_2, \{0,1\}, \delta_2, p_0, F_2)$, kde $K_2 = \{p_0, p_1, p_2, p_3, p_4, p_5\}, F_2 = \{p_5\}$ a δ_2 je daná podľa obrázka:



Napíšme si hypotézy pre triedy zodpovedajúce stavom:

$$
\begin{aligned}
Kl[p_0] &= \{\lambda\} \\
Kl[p_1] &= \{1\} \\
Kl[p_2] &= \{0w, w \in \{0,1\}^*\} \cup \{11w, w \in \{0,1\}^*\} \\
Kl[p_3] &= \{10w, w \in \{1,01\}^*\} \\
Kl[p_4] &= \{10w0, w \in \{1,01\}^*\} \\
Kl[p_5] &= \{10w00y, w \in \{1,01\}^*, y \in \{0,1\}^*\}
\end{aligned}
$$

Ľahko by sa ukázalo, že triedy sú disjunktné a dokopy dávajú celú Σ^*. Tento fakt ale nie sme povinní priamo dokazovať (aj keď triedy musia mať túto vlastnosť). Pri dokazovaní správnosti popisu tried totiž dokážeme nepriamo aj tento bočný fakt. Dôkaz správnosti našej hypotézy spravíme opäť indukciou na dĺžku vstupného slova.

Vyriešme ale pre jednoduchosť najprv jeden špeciálny prípad: triedu $Kl[p_0]$. Slovo λ do tejto triedy patrí (ako pre každý počiatočný stav každého automatu). Do tejto triedy ale nepatrí žiadne iné slovo: dá sa ľahko ukázať, že žiadny výpočet sa do stavu p_0 nevráti. Ak by totiž niekedy počas výpočtu automat do p_0 vrátil, musel by v tomto výpočte byť použitý krok $\delta_2(p, a) = p_0$ pre nejaké $p \in K_2, a \in \Sigma$. Lenže z definície δ_2 vidíme, že toto sa nemôže stať (slovne povedané, ťažko sa do p_0 vrátime, keď do neho nesmeruje žiadna šipka.)

Báza indukcie: Ak $w = \lambda$, výpočet skončí v stave p_0. Ak $|w| = 1$, tak ak $w = 0$, podľa hypotézy $w \in Kl[p_2]$ a výpočet skončí v p_2. Ak $w = 1$, potom $w \in Kl[p_1]$ a výpočet skončí v p_1. Všetky tri možnosti teda zodpovedajú našej hypotéze.

Indukčný krok: Ak $|w| = i > 1$, potom $w = ua$, kde $u \in \{0,1\}^{i-1}, a \in \{0,1\}$. Výpočet na w sa teda skladá z výpočtu na u a poslednom kroku na a:

$(q_0, ua) \vdash^{i-1} (\hat{\delta}(q_0, u), a) \vdash (\delta(\hat{\delta}(q_0, u), a), \lambda)$.

Keďže $|u| < i$, potom sa na výpočet na u vzťahuje indukčný predpoklad, takže vzhľadom na stav, v ktorom skončil výpočet na u môžeme predpokladať tvar slova u podľa príslušnej triedy. Keďže máme dokázanú triedu $Kl[p_0]$, vidíme, že v tejto triede nie je žiadne slovo dĺžky $i - 1 \geq 1$. Preto budeme uvažovať len prípady, kedy výpočet na prefixe u skončil v niektorom z ostatných stavov. Uvažujme, ake prípady môžu nastať:

- $\hat{\delta}_2(p_0, u) = p_1$, $a = 0$, potom $w = 10$, takže w má patriť do $Kl[p_3]$. Keďže $\delta_2(p_1, 0) = p_3$, potom nám tento výpočet neporuší hypotézu.

- $\hat{\delta_2}(p_0, u) = p_1$, $a = 1$, potom $w = 11$, takže w má patriť do $Kl[p_2]$. Keďže $\delta_2(p_1, 1) = p_2$, potom nám tento výpočet neporuší hypotézu.
- $\hat{\delta_2}(p_0, u) = p_2$. Potom $u = 0y$ alebo $u = 11y$, kde $y \in \{0,1\}^*$. Či už za u pridáme 0 alebo 1, w má patriť do $Kl[p_3]$. Keďže $\delta_2(p_2, 0) = \delta_2(p_2, 1) = p_2$, potom nám tento výpočet neporuší hypotézu.
- $\hat{\delta_2}(p_0, u) = p_3$, potom $u = 10y$, pre nejaké $y \in \{1, 01\}^*$. Potom ak $a = 0$, $w = 10y0$, takže vidíme, že $w \in Kl[p_4]$, čomu zodpovedá aj funkcia δ_2. Ak $a = 1$, $w = 10y1$, takže ak $y \in \{1, 01\}^*$, potom aj $y1 \in \{1, 01\}^*$, a teda $w \in Kl[p_3]$. Keďže $\delta_2(p_3, 1) = p_3$, potom nám tento výpočet neporuší hypotézu.
- $\hat{\delta_2}(p_0, u) = p_4$, potom $u = 10y0$ pre nejaké $y \in \{1, 01\}^*$. Ak $a = 0$, potom $w = 10y00$, takže w má patriť do $Kl[p_5]$. Ak $a = 1$, potom $w = 10y01$, takže ak $y \in \{1, 01\}^*$, potom aj $y01 \in \{1, 01\}^*$, takže w má patriť do $Kl[p_3]$. Obom prípadom zodpovedá aj prechodová funkcia automatu M_2.
- $\hat{\delta_2}(p_0, u) = p_5$, potom $u = 10y00z$, kde $y \in \{1, 01\}^*, z \in \{0, 1\}^*$. Akékoľvek je a, slovo $za \in \{0, 1\}^*$, čiže w má patriť do $Kl[p_5]$. Pre oba prípady je ale $\delta_2(p_5, a) = p_5$, takže aj táto možnosť je v poriadku.

Rozobrali sme všetky možnosti a dokázali sme platnosť tried aj pre slová s dĺžkou i, čím sme dokončili dôkaz správnosti tried. Keďže $F_2 = \{p_5\}$, potom $L(M_2) = Kl[p_5] = \{w = 10y00z | y \in \{1, 01\}^*, z \in \{0, 1\}^*\} = \{w = 10y00z | y, z \in \{0, 1\}^*\} = L_2$. Rovnosť $\{w = 10y00z | y \in \{1, 01\}^*, z \in \{0, 1\}^*\} = \{w = 10y00z | y, z \in \{0, 1\}^*\}$ sa dokáže ľahko. Stačí si uvedomiť, že množina naľavo je priamo z množinového zápisu podmnožinou tej napravo. Naopak, ak máme slovo, ktoré patrí napravo, potom sa dá jednoducho napísať ako slovo patriace naľavo - slovo prepíšeme podľa prvého výskytu podslova 00. Tento dôkaz síce nie je plnohodnotný formálny, avšak v tejto úlohe máme dokazovať hlavne iné veci.

Takže máme pripravené dve súčiastky výsledného automatu a môžeme skonštruovať taký M, že $L(M) = L(M_1) \cap L(M_2) = L_1 \cap L_2 = L$. Všeobecná metóda simulácie pre prienik dvoch automatov nás privedie k $M = (K_1 \times K_2, \{0, 1\}, \delta_M, (q_0, p_0), F_M)$, kde $\delta_M((q_i, p_j), a) = (\delta_1(q_i, a), \delta_2(p_j, a))$. Slovne povedané, stavy nového automatu M budu dvojice stavov pôvodného, pričom M si bude v týchto stavoch pamätať stavy pôvodných automatov na vstupnom slove. M má akceptovať len ak akceptovali pôvodné autoamty, preto $F_M = F_1 \times F_2 = \{(q_0, p_5)\}$.

Bodovanie + nedostatky Za rozpísanie si L na prienik dvoch jazykov sa udeľoval jeden bod. Za automat pre paritu jednotiek boli tri body: jeden za konštrukciu a dva za poriadny dôkaz správnosti. Za druhý automat, keďže bol zložitejší, boli až 4 body - jeden za konštrukciu a tri za dôkaz. Zvyšné dva body boli za výsledný automat. Udeľovali sa podľa kvality realizácie. Ak ste bez slova nakreslili hromadu kruhov a pospájali ich šipkami, tak ste dva body nedostali. Na dva body stačilo dvoma vetami (prípadne pár formálnymi vzťahmi) vysvetliť, ako ste tento automat získali, ako ste prišli na jeho prechodovú funkciu a ako ste určili akceptačné stavy.

Automaty ste mali skoro vždy správne, väčšina strhnutých bodov bola za chýbajúce alebo nedostatočné dôkazy. Ako mali dôkaz vyzerať, si pozrite vyššie. Medzi všeobecnými komentármi by som uviedol: keď definujeme triedy $KL[q]$, tak musia byť disjunktné (mať prázdny prienik) a všetky triedy dokopy musia dať Σ^*. Uvedomme si, že trieda $KL[q]$ je množina slov, ktorých výpočet končí v stave q. Výpočet na každom slove skončí v nejakom stave a žiadne slovo neskončí v dvoch stavoch naraz. Okrem tohoto ste niektorí nesprávne rozdelili jazyk na prienik dvoch iných. Tým nechcem povedať, že sa to dalo len jedným spôsobom. Akceptoval by som ľubovoľné správne riešenie.

Vzorové riešenia úlohy 5.2

V mnou vytvorenom konečnom automate budem používať myšlienku uloženia si informácie do stavov automatu. Konkrétne si budem pamätať posledné 3 prečítané znaky.

$$\beta(w) = \begin{cases} w & \text{if } |w| < 3 \\ a_{n-2}a_{n-1}a_n & \text{if } |w| \geq 3 \wedge w = a_1 \dots a_n \end{cases} \quad (1)$$

$$\begin{aligned}
M &= (Q, \Sigma, \delta, q_\lambda, F) && (2) \\
Q &= \{q_w \mid w \in \Sigma^* \wedge |w| \leq 3\} \bigcup \{q_f\} && (3) \\
F &= \{q_f\} && (4) \\
\delta &: (Q \times \Sigma) \to Q && (5) \\
\delta(q_w, c) &= \begin{cases} q_{\beta(w\cdot c)} & \text{if } w \cdot c \notin \{aaaa, aaba, abaa, abba\} \\ q_f & \text{if } w \cdot c \in \{aaaa, aaba, abaa, abba\} \end{cases} && (6) \\
\delta(q_f, c) &= q_f && (7)
\end{aligned}$$

$L(M) \subseteq L$:

$$\begin{aligned}
w &= c_0c_1c_2 \dots c_n && (8) \\
w \in L(M) &\Rightarrow (q_\lambda, c_0 \dots c_n) \\
&\vdash_M (q_{v_1}, c_1 \dots c_n) \\
&\vdash_M (q_{v_2}, c_2 \dots c_n) \\
&\vdash_M \dots \vdash_M (q_{v_i}, c_i \dots c_n) \\
&\vdash_M (q_f, c_{i+1} \dots c_n) \\
&\vdash_M (q_f, c_{i+2} \dots c_n) \\
&\vdash_M \dots \vdash_M (q_f, \lambda) && (9) \\
(q_{v_i}, c_i \dots c_n) \vdash_M (q_f, c_{i+1} \dots c_n) &\Rightarrow v_i \cdot c_i \in \{aaaa, aaba, abaa, abba\} && (10) \\
&\Rightarrow w = c_0 \dots c_{i-4} \cdot x \cdot c_{i+1} \dots c_n \\
&\wedge x \in \{aaaa, aaba, abaa, abba\} && (11) \\
&\Rightarrow w \in L && (12)
\end{aligned}$$

$L(M) \supseteq L$:

Podľa definície L každé slovo jazyka L musí obsahovať podslovo z množiny $X = \{aaaa, aaba, abaa, abba\}$. Z konštrukcie automatu M je zrejmé, že automat v 15tich zo svojich stavov rozlišuje posledné prečítané podslovo maximálnej dĺžky 3. Počas čítania slova $w \in L$ automat musí prečítať niektoré z podslov X. Podľa definície sa automat v takom prípade dostane do stavu q_f, v ktorom zotrvá až do konca výpočtu akceptácie slova a slovo následne akceptuje.

$L(M) = L$:

$$L(M) \subseteq L \wedge L(M) \supseteq L \Rightarrow L(M) = L \quad (13)$$

Časté chyby: Najčastejšie chyby boli v dôkaze indukciou. Treba si vždy uvedomiť, čo chceme dokázať, teda ako má vyzerať indukčný prepoklad resp. tvrdenie $P(n)$, z ktorého dokazujeme $P(n+1)$. Chceme dokázať, že $w \in L(M) \Leftrightarrow w \in L$. Teda $P(n)$ je toto tvrdenie pre $|w| \leq n$ resp. $|w| = n$. Mnohí ste mali indukčný predpoklad dodržanie triedy podľa stavov. A ukazovali ste, že keď $w \in KL[q_i]$, tak $wa \in KL[q_j]$ resp. $wb \in KL[q_k]$, na základe prechodovej funkcie. Treba si uvedmiť, že to je len prepis prechodovej funkcie.

PS: Za poskytnutie riešenia ďakujem Petrovi Pokojnému.

Vzorové riešenie úlohy 5.3

Časť a) Predpokladajme, že existuje KA akceptujúci L, ktorý má menej ako 10 stavov. Z dirichletovho princípu vyplýva, že existujú dve slová z L, a^i a a^j, kde $0 \leq i < j \leq 9$ také, že končia obe v stave q. Potom ale vieme, že slová $a^i a^{9-i} = a^9$ a $a^j a^{9-i} = a^{9+k}$ $(k > 1)$ budú končiť v rovnakom stave, pritom prvé slovo má byť akceptované a druhé slovo nemá byť akceptované. Z toho sme dostali spor s tým, že KA mal menej ako 10 stavov.

Časť b) Tento jazyk bol regulárny, aj keď sa na prvý pohľad mohlo zdať, že regulárny nie je. Ukážeme, že $L = \{xy \mid |x|_a = |y|_b\} = \{w \mid |w|_a = 2k, k \geq 0\} = L'$. Rovnosť jazykov ukážeme pomocou dvoch inklúcií.

Najprv ukážeme, že $L \subseteq L'$. Majme slovo, ktoré sa dá rozdeliť na dve podslová obsahujúce rovnaký počet písmen a. Potom celé toto slovo musí mať párny počet písmen a, pretože z $|x|_a = |y|_a = k$ vypýva, že $|xy|_a = 2k$.

Opačne, chceme ukázať, že $L' \subseteq L$, teda každé slovo s párnym počtom a sa dá rozdeliť na dve slová s rovnakým počtom a. To môžeme ukázať napríklad nasledovnou úvahou. Ak rozdelíme slovo $w = w_1 w_2 \ldots w_n$ na slová $x_0 = \lambda$ a $y_0 = w$, vidíme, že $|x_1|_a = 0$ a $|y_1|_a = 2k$. Ak ho rozdelíme na $x_n = w$ a $y_n = \lambda$, dostávame opačný pomer, $|x_n|_a = 2k$ a $|y_n|_a = 0$. Postupne vieme zadefinovať všetky možné rozdelenia, ako $x_i = w_1 \ldots w_i$, $y_i = w_{i+1} \ldots w_n$. Ľahko ukážeme, že môžu nastať dva prípady: buď $|x_i|_a = |x_{i-1}|_a$ alebo $|x_i|_a = |x_{i-1}|_a + 2$. Ak vieme, že $|x_0|_a = 0$ a $|x_n|_a = 2k$, zmena z $|x_{i-1}|_a$ na $|x_i|_a$ je buď o 2 alebo o 0, ľahko vidíme, že existuje isté ℓ také, že $|x_\ell|_a = k = |y_\ell|_a$.

Teraz už len stačí nájsť automat pre L' a dokázať, že generuje L'. Tento automat M bude mať dva stavy, q_0 a q_1 na písmeno b sa stav meniť nebude, na písmeno a sa stav zmení na opačný. Akceptačným stavom bude stav q_0. Ukážeme, že každé slovo z L' bude akceptované týmto automatom, a naopak, že každé slovo, ktoré je akceptované automatom, je z L'. Majme slovo $w \in L'$, toto slovo obsahuje $2k$ áčok. Písmená b nás nezaujímajú, pretože KA nemení svoj stav, keď číta b. Preto sa stav zmení presne $2k$ krát, čo znamená, že skončíme v stave q_0 a teda slovo w akceptujeme. Naopak, ak máme slovo $u \in L(M)$, vieme, že akceptačný stav je len jeden, q_0.

Indukciou vzhľadom na dĺžku slova ukážeme, že ak slovo končí v stave q_0, má párny počet písmen a a ak v stave q_1 tak má nepárny počet písmen a. Pre $|u| = 0$, $u = \lambda$ to platí, automat skončí v stave q_0. Nech to platí pre ľubovoľné slovo dĺžky $n - 1$. Chceme ukázať, že to platí ak pre slovo u dĺžky n. Slovo u sa dá napísať ako $u = vx$, kde v je slovo dlžky $n - 1$ a x je jeden znak. Rozoberieme prípad, keď u má párny počet písmen a, druhý prípad by sa rozobral analogicky.

Ak je $x = a$, vieme, že v má nepárny počet a a teda končí v stave q_1. Po prečítaní a sa dostávame do stavu q_0, preto u končí v q_0 a u má párny počet písmen a.

Ak je $x = b$, vieme, že v má párny počet a a teda končí v stave q_0. Po prečítaní b sa stav nemení, preto aj u končí v q_0 a u má párny počet písmen a.