Poznámky k cvičeniu č. 5

Peter Kostolányi

19. októbra 2016

Konečné automaty

V nasledujúcom iba stručne zopakujeme definície deterministických a nedeterministických konečných automatov, ktoré boli zavedené na prednáške.

Definícia 1. *Deterministický konečný automat* (DKA) je pätica $A = (K, \Sigma, \delta, q_0, F)$, kde K je neprázdna konečná množina stavov, Σ je vstupná abeceda, $\delta\colon K \times \Sigma \to K$ je prechodová funkcia, $q_0 \in K$ je počiatočný stav a $F \subseteq K$ je množina akceptačných stavov.

Definícia 2. Nech $A = (K, \Sigma, \delta, q_0, F)$ je deterministický konečný automat. *Konfigurácia* automatu A je usporiadaná dvojica (q, w), kde $q \in K$ je stav a $w \in \Sigma^*$ je slovo (reprezentujúce nedočítanú časť vstupného slova).

Definícia 3. Nech $A = (K, \Sigma, \delta, q_0, F)$ je deterministický konečný automat. *Krok výpočtu* automatu A je binárna relácia $\vdash_A$ na množine konfigurácií automatu A taká, že pre všetky dvojice stavov $p, q \in K$, symboly $c \in \Sigma$ a slová $w \in \Sigma^*$ platí $(p, cw) \vdash_A (q, w)$ práve vtedy, keď $\delta(p, c) = q$. Ak je automat A zrejmý z kontextu, píšeme namiesto $\vdash_A$ iba $\vdash$.

Krok výpočtu konečného automatu je definovaný ako relácia, a teda pre každé $k \in \mathbb{N}$ sú v relácii $\vdash^k$ tie dvojice konfigurácií, medzi ktorými možno „prejsť“ na k krokov výpočtu. V relácii $\vdash^*$, reflexívno-tranzitívnom uzávere relácie $\vdash$, sú tie dvojice konfigurácií, medzi ktorými možno „prejsť“ na nejaký počet krokov a v tranzitívnom uzávere, relácii $\vdash^+$, sú tie dvojice konfigurácií, medzi ktorými možno „prejsť“ na nejaký nenulový počet krokov. Relácia $\vdash^0$ je identita.

Podobne ako pri odvodeniach v bezkontextových gramatikách, aj tu je zvykom zneužívať notáciu a hovoriť o „výpočte $(q_0, a_1 a_2 \ldots a_n) \vdash (q_1, a_2 \ldots a_n) \vdash^* (q_n, \varepsilon)$“ a podobne, aj keď formálne nejde o označenie žiadneho matematického objektu. Na rozdiel od odvodení v gramatikách však takýto zápis aspoň nesie úplnú informáciu (ak by sme ho neskrátili použitím uzáveru a vypísali by sme všetky konfigurácie). Inými slovami, výpočet konečného automatu by bolo možné definovať ako postupnosť konfigurácií, kde každé dve po sebe idúce sú v relácii $\vdash$.

Definícia 4. Nech $A = (K, \Sigma, \delta, q_0, F)$ je deterministický konečný automat. *Jazyk* akceptovaný automatom A je definovaný ako $L(A) = \{w \in \Sigma^* \mid \exists q \in F : (q_0, w) \vdash^* (q, \varepsilon)\}$.

Poznámka 1. V literatúre sa deterministické konečné automaty občas zvyknú definovať s čiastočnou prechodovou funkciou, čo znamená, že hodnota $\delta(q, c)$ nemusí byť definovaná pre všetky stavy q a symboly c.[1] Na tomto predmete však pracujeme s definíciou s úplnou prechodovou funkciou. V úlohách na zostrojenie deterministických konečných automatov je preto dôležité dbať na to, aby bola prechodová funkcia naozaj „všade“ definovaná.

Nedeterministický konečný automat sa oproti deterministickému líši tým, že pre jeden stav a symbol môže existovať aj viac výstupov prechodovej funkcie (prípadne nemusí existovať žiaden). Formálne tak prechodová funkcia vracia pre každý stav q a symbol c *množinu* stavov $\delta(q, c)$. Stavy z tejto množiny intuitívne zodpovedajú možnostiam pre „nedeterministické rozhodnutia“ konečného automatu: v každom kroku výpočtu si automat môže „vybrať“ jeden stav zo zodpovedajúcej množiny, do ktorého prejde. Automat potom akceptuje nejaké slovo w v prípade, že *existuje* postupnosť „nedeterministických rozhodnutí“, pomocou ktorých dočíta slovo w v akceptačnom stave. Nedeterministické konečné automaty navyše majú dovolené prechody, pri ktorých nič neprečítajú zo vstupu, tzv. „prechody na ε“.

[1] Ekvivalencia oboch definícií je priamym dôsledkom ekvivalencie deterministických a nedeterministických konečných automatov, ktorej sa budeme venovať na nasledujúcom cvičení.

Definícia 5. *Nedeterministický konečný automat* (NKA) je pätica $A = (K, \Sigma, \delta, q_0, F)$, kde K je neprázdna konečná množina stavov, Σ je vstupná abeceda, $\delta\colon K \times (\Sigma \cup \{\varepsilon\}) \to 2^K$ je prechodová funkcia, $q_0 \in K$ je počiatočný stav a $F \subseteq K$ je množina akceptačných stavov.

Definícia 6. Nech $A = (K, \Sigma, \delta, q_0, F)$ je nedeterministický konečný automat. *Konfigurácia* automatu A je usporiadaná dvojica (q, w), kde $q \in K$ je stav a $w \in \Sigma^*$ je slovo (reprezentujúce nedočítanú časť vstupného slova).

Definícia 7. Nech $A = (K, \Sigma, \delta, q_0, F)$ je nedeterministický konečný automat. *Krok výpočtu* automatu A je binárna relácia $\vdash_A$ na množine konfigurácií automatu A taká, že pre všetky dvojice stavov $p, q \in K$, slová $z \in \Sigma \cup \{\varepsilon\}$ a slová $w \in \Sigma^*$ platí $(p, zw) \vdash_A (q, w)$ práve vtedy, keď $q \in \delta(p, z)$. Ak je automat A zrejmý z kontextu, píšeme namiesto $\vdash_A$ iba $\vdash$.

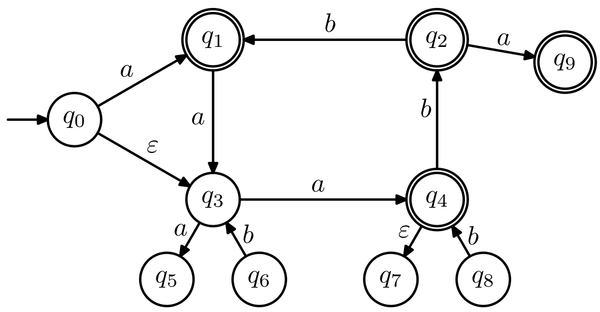
Definícia 8. Nech $A = (K, \Sigma, \delta, q_0, F)$ je nedeterministický konečný automat. *Jazyk* akceptovaný automatom A je definovaný ako $L(A) = \{w \in \Sigma^* \mid \exists q \in F : (q_0, w) \vdash^* (q, \varepsilon)\}$.

Častým spôsobom ako zadať konečný automat je tzv. *prechodový diagram*, v ktorom sú stavy znázornené „kolečkami“, prechodová funkcia šípkami medzi stavmi, počiatočný stav krátkou šípkou do zodpovedajúceho „kolečka“ a akceptačné stavy „zdvojenými kolečkami“.

Príklad 1. Uvažujme nedeterministický konečný automat $A = (K, \Sigma, \delta, q_0, F)$ s $K = \{q_0, \ldots, q_9\}$, $\Sigma = \{a, b\}$, $F = \{q_1, q_2, q_4, q_9\}$ a prechodovou funkciou danou nasledovne (uvádzame iba neprázdne množiny $\delta(q, z)$):

$$\begin{aligned} \delta(q_0, a) &= \{q_1\}, & \delta(q_3, a) &= \{q_4, q_5\}, \\ \delta(q_0, \varepsilon) &= \{q_3\}, & \delta(q_4, b) &= \{q_2\}, \\ \delta(q_1, a) &= \{q_3\}, & \delta(q_4, \varepsilon) &= \{q_7\}, \\ \delta(q_2, a) &= \{q_9\}, & \delta(q_6, b) &= \{q_3\}, \\ \delta(q_2, b) &= \{q_1\}, & \delta(q_8, b) &= \{q_4\}. \end{aligned}$$

Prechodový diagram automatu A je na obrázku 1.



Obr. 1: Prechodový diagram nedeterministického konečného automatu A. Automat A **nie je** v „prasiatkovom“ normálnom tvare.

Nasledujúce jednoduché tvrdenie, ktorého dôkaz prenechávame na domácu úlohu, budeme často používať bez toho, aby sme to explicitne uvádzali (platí pre deterministické aj pre nedeterministické konečné automaty):

Tvrdenie 1. *Nech $A = (K, \Sigma, \delta, q_0, F)$ je konečný automat. Potom platí:*

a) $\forall p, q \in K\ \forall u, v \in \Sigma^* : (p, uv) \vdash^* (q, v) \iff (p, u) \vdash^* (q, \varepsilon)$,

b) $\forall p, q, r \in K\ \forall u, v \in \Sigma^* : (p, u) \vdash^* (q, \varepsilon) \ \wedge\ (q, v) \vdash^* (r, \varepsilon) \Rightarrow (p, uv) \vdash^* (r, \varepsilon)$.

Dokazovanie tvrdení typu $L(A) = L$

Dokazovanie tvrdenia $L(A) = L$ pre konečný automat A a jazyk L je v mnohom podobné dokazovaniu takýchto tvrdení pre bezkontextové gramatiky. Hlavným spoločným znakom je matematická indukcia, ktorou sa obyčajne dokáže o niečo silnejšie tvrdenie, než to pôvodne zamýšľané. Pre konečné automaty sú takýmto vhodným silnejším tvrdením väčšinou *invarianty* pre jednotlivé stavy. Pre každý stav q sa teda presne charakterizujú slová w, pre ktoré existuje výpočet $(q_0, w) \vdash^* (q, \varepsilon)$. Tvrdenie $L(A) = L$ potom vyplýva z invariantov pre akceptačné stavy.

Úloha 1. Zostrojte deterministický konečný automat akceptujúci jazyk

$$L = \{w \in \{a,b\}^* \mid \#_a(w) \equiv 1 \pmod 3 \ \wedge\ \#_b(w) \equiv 3 \pmod 7\}$$

a správnosť svojej konštrukcie dokážte.

Riešenie. Pri konštrukcii automatu akceptujúceho jazyk L na chvíľu upustíme od zvyčajnej notačnej konvencie označovania stavov symbolmi p, q s prípadnými „ozdobami“ a za množinu stavov vezmeme karteziánsky súčin $\mathbb{Z}_3 \times \mathbb{Z}_7$. Čitateľ nespokojný s takýmto prístupom iste dokáže konštrukciu zapísať aj s rešpektovaním bežných konvencií (a bez prílišnej straty na elegancii konštrukcie).

Automat $A = (K, \Sigma, \delta, q_0, F)$ bude mať $K = \mathbb{Z}_3 \times \mathbb{Z}_7$, $\Sigma = \{a, b\}$, $q_0 = [0, 0]$, $F = \{[1, 3]\}$ a

$$\delta([s,t], a) = [s+1, t], \quad \forall s \in \mathbb{Z}_3\ \forall t \in \mathbb{Z}_7,$$
$$\delta([s,t], b) = [s, t+1], \quad \forall s \in \mathbb{Z}_3\ \forall t \in \mathbb{Z}_7,$$

kde operácia sčítania je vždy modulo príslušné číslo.

Namiesto tvrdenia $L(A) = L$ dokážeme najprv silnejšie tvrdenie – invarianty pre jednotlivé stavy – z ktorého neskôr správnosť našej konštrukcie vyplynie:

$$\forall [s,t] \in K\ \forall w \in \Sigma^* : ([0,0], w) \vdash^* ([s,t], \varepsilon) \iff \#_a(w) \equiv s \pmod 3 \ \wedge\ \#_b(w) \equiv t \pmod 7.$$

Dokážme teraz jednotlivé implikácie, vždy naraz pre všetky slová w a stavy $[s, t]$:

$\Rightarrow$: Matematickou indukciou vzhľadom na dĺžku výpočtu n dokážeme, že pre všetky slová $w \in \Sigma^*$ a všetky stavy $[s, t] \in K$ platí

„ak $([0,0], w) \vdash^n ([s,t], \varepsilon)$, tak $\#_a(w) \equiv s \pmod 3 \ \wedge\ \#_b(w) \equiv t \pmod 7$“.

1. Nech $n = 0$. Potom $[s, t] = [0, 0]$ a $w = \varepsilon$. Tvrdenie teda platí.
2. Nech tvrdenie platí pre $n = k$. Ukážeme, že platí aj pre $n = k + 1$.
   Nech $w \in \Sigma^*$ je slovo, nech $[s, t] \in K$ je stav a nech platí $([0,0], w) \vdash^{k+1} ([s,t], \varepsilon)$. To možno rozpísať ako $([0,0], uc) \vdash^k ([s',t'], c) \vdash ([s,t], \varepsilon)$, kde $[s', t'] \in K$, $u \in \Sigma^*$, $c \in \Sigma$ (keďže automat A je deterministický) a $w = uc$.
   a) Nech $c = a$. Potom pre stav $[s', t']$ musí platiť $\delta([s',t'], a) = [s, t]$. Z definície prechodovej funkcie δ vyplýva, že $s' = s - 1$ a $t' = t$ a z indukčného predpokladu ďalej plynie $\#_a(u) \equiv s - 1 \pmod 3$ a $\#_b(u) \equiv t \pmod 7$. To už priamo implikuje $\#_a(w) = \#_a(ua) = \#_a(u) + \#_a(a) \equiv s \pmod 3$ a $\#_b(w) = \#_b(ua) = \#_b(u) \equiv t \pmod 7$, čo bolo treba dokázať.
   b) Prípad $c = b$ je analogický.

$\Leftarrow$: Keďže je automat A deterministický, je možné dokázať opačnú implikáciu aj bez ďalšej indukcie. Stačí si všimnúť, že dokazované invarianty pre jednotlivé stavy sú „disjunktné“ (nemôže sa teda stať, že by pre nejaké slovo súčasne platili invarianty pre dva rôzne stavy). Z determinizmu automatu A ďalej vyplýva, že každé slovo w dočíta až do konca. Nech pre takéto slovo platí invariant pre nejaký stav p. Ak automat A dočítal slovo w v nejakom stave $q \neq p$, musel by podľa predchádzajúcej implikácie platiť aj invariant pre q, čo sa vylučuje s platnosťou invariantu pre p. Automat A teda musí dočítať slovo w v stave p.

Dôkaz samotného tvrdenia $L(A) = L$ je už možné urobiť jednoducho:

$\subseteq$: Nech $w \in L(A)$. Keďže $[1,3]$ je jediný akceptačný stav, musí platiť $([0,0], w) \vdash^* ([1,3], \varepsilon)$. Z invariantu pre $[1,3]$ ale vyplýva $\#_a(w) \equiv 1 \pmod 3$ a $\#_b(w) \equiv 3 \pmod 7$, a teda $w \in L$.

$\supseteq$: Nech $w \in L$. Potom platí $\#_a(w) \equiv 1 \pmod 3$ a $\#_b(w) \equiv 3 \pmod 7$ a z dokázaných invariantov vyplýva, že $([0,0], w) \vdash^* ([1,3], \varepsilon)$. Preto $w \in L(A)$.

Tvrdenie je dokázané. $\square$

Úloha 2. Zostrojte nedeterministický konečný automat akceptujúci jazyk

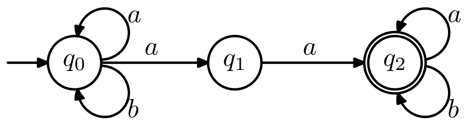
$$L = \{uaav \mid u, v \in \{a,b\}^*\}$$

a správnosť svojej konštrukcie dokážte.

Riešenie. Automat $A = (K, \Sigma, \delta, q_0, F)$ akceptujúci L zostrojíme nasledovne: $K = \{q_0, q_1, q_2\}$, $\Sigma = \{a, b\}$, $F = \{q_2\}$ a

$$\begin{aligned}
&\delta(q_0, a) = \{q_0, q_1\}, \qquad &&\delta(q_1, a) = \{q_2\}, \qquad &&\delta(q_2, a) = \{q_2\},\\
&\delta(q_0, b) = \{q_0\}, &&\delta(q_1, b) = \emptyset, &&\delta(q_2, b) = \{q_2\}.
\end{aligned}$$

Prechodový diagram automatu A je na obrázku 2. Pre stavy automatu A dokážeme nasledujúcu



Obr. 2: Prechodový diagram nedeterministického konečného automatu A.

sadu invariantov:

$$\begin{aligned}
&I_0 \colon \forall w \in \Sigma^* : (q_0, w) \vdash^* (q_0, \varepsilon) \iff w \in \Sigma^*,\\
&I_1 \colon \forall w \in \Sigma^* : (q_0, w) \vdash^* (q_1, \varepsilon) \iff w \in \Sigma^* \cdot \{a\},\\
&I_2 \colon \forall w \in \Sigma^* : (q_0, w) \vdash^* (q_2, \varepsilon) \iff w \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*.
\end{aligned}$$

Dokážeme postupne obidve implikácie, zakaždým pre všetky tri invarianty naraz.

$\Rightarrow$: Matematickou indukciou vzhľadom na n dokážeme nasledujúce implikácie:

$$\begin{aligned}
&I'_0 \colon \forall w \in \Sigma^* : (q_0, w) \vdash^n (q_0, \varepsilon) \Rightarrow w \in \Sigma^*,\\
&I'_1 \colon \forall w \in \Sigma^* : (q_0, w) \vdash^n (q_1, \varepsilon) \Rightarrow w \in \Sigma^* \cdot \{a\},\\
&I'_2 \colon \forall w \in \Sigma^* : (q_0, w) \vdash^n (q_2, \varepsilon) \Rightarrow w \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*.
\end{aligned}$$

1. Pre $n = 0$ môže byť antecedent[2] splnený iba pre implikáciu I'_0. V takom prípade tiež nutne $w = \varepsilon$, čo je slovo zo Σ^* a tvrdenie teda platí.
2. Nech implikácie I'_0 až I'_2 platia pre $n = k$. Ukážeme, že platia aj pre $n = k + 1$, zvlášť pre každú z nich:

   I'_0: Nech $w \in \Sigma^*$ je slovo také, že $(q_0, w) \vdash^{k+1} (q_0, \varepsilon)$. Potom triviálne $w \in \Sigma^*$ a implikácia platí.

   I'_1: Nech $w \in \Sigma^*$ je slovo také, že $(q_0, w) \vdash^{k+1} (q_1, \varepsilon)$. To možno rozpísať aj ako $(q_0, uz) \vdash^k (q, z) \vdash (q_1, \varepsilon)$, kde $q \in K$, $u \in \Sigma^*$, $z \in \Sigma \cup \{\varepsilon\}$ a $w = uz$.

   a) Nech $z = a$. Potom $w = ua \in \Sigma^* \cdot \{a\}$ a implikácia platí.

[2]Ľavá strana implikácie.

b) Nech $z = b$. Pre stav q musí platiť $q_1 \in \delta(q, b)$. Takýto stav ale neexistuje, takže implikácia triviálne platí.

c) Prípad $z = \varepsilon$ je analogický prípadu $z = b$.

I_2': Nech $w \in \Sigma^*$ je slovo také, že $(q_0, w) \vdash^{k+1} (q_2, \varepsilon)$. To možno rozpísať aj ako $(q_0, uz) \vdash^k (q, z) \vdash (q_2, \varepsilon)$, kde $q \in K$, $u \in \Sigma^*$, $z \in \Sigma \cup \{\varepsilon\}$ a $w = uz$.

a) Nech $z = a$. Pre stav q musí platiť $q_2 \in \delta(q, a)$, a teda $q \in \{q_1, q_2\}$. Ak $q = q_1$, z indukčného predpokladu na implikáciu I_1' vyplýva $u \in \Sigma^* \cdot \{a\}$. Preto $w = ua \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$ a implikácia platí. Ak $q = q_2$, z indukčného predpokladu na implikáciu I_2' vyplýva $u \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$, a teda aj $w = ua \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$.

b) Nech $z = b$. Pre stav q musí platiť $q_2 \in \delta(q, b)$, a teda $q = q_2$. Z indukčného predpokladu na implikáciu I_2' máme $u \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$, z čoho už priamo vyplýva $w = ub \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$.

c) Nech $z = \varepsilon$. Pre stav q musí platiť $q_2 \in \delta(q, \varepsilon)$. Takýto stav q ale neexistuje a implikácia je teda triviálne splnená.

$\Leftarrow$: Implikáciu sprava doľava dokážeme postupne pre jednotlivé invarianty. Najprv ukážeme, že platí implikácia

$$I_0'' \colon w \in \Sigma^* \Rightarrow (q_0, w) \vdash^* (q_0, \varepsilon).$$

Indukciou vzhľadom na dĺžku slova w.

1. Nech $|w| = 0$. Potom $w = \varepsilon$ a triviálne platí $(q_0, \varepsilon) \vdash^* (q_0, \varepsilon)$.
2. Nech implikácia platí pre všetky w s $|w| = k$. Ukážeme, že platí aj pre všetky w s $|w| = k + 1$.
   Nech $w \in \Sigma^{k+1}$. Potom $w = uc$, kde $u \in \Sigma^k$ a $c \in \Sigma$. Na slovo u sa vzťahuje indukčný predpoklad, a teda $(q_0, u) \vdash^* (q_0, \varepsilon)$. Platí však $q_0 \in \delta(q_0, a)$, ako aj $q_0 \in \delta(q_0, b)$. Preto $q_0 \in \delta(q_0, c)$ a $(q_0, c) \vdash (q_0, \varepsilon)$. Výpočet existujúci podľa indukčného predpokladu teda môžeme predĺžiť na $(q_0, uc) \vdash^* (q_0, c) \vdash (q_0, \varepsilon)$ a implikácia platí.

Dokážeme teraz druhú implikáciu

$$I_1'' \colon w \in \Sigma^* \cdot \{a\} \Rightarrow (q_0, w) \vdash^* (q_1, \varepsilon).$$

To ide aj bez indukcie. Každé slovo $w \in \Sigma^* \cdot \{a\}$ totiž možno napísať ako $w = ua$, kde $u \in \Sigma^*$. Podľa už dokázanej implikácie I_0'' platí $(q_0, u) \vdash^* (q_0, \varepsilon)$. Keďže ale $q_1 \in \delta(q_0, a)$, dostávame $(q_0, ua) \vdash^* (q_0, a) \vdash (q_1, \varepsilon)$, čo bolo treba dokázať.

Dokážme teda poslednú implikáciu

$$I_2'' \colon w \in \Sigma^* \cdot \{aa\} \cdot \Sigma^* \Rightarrow (q_0, w) \vdash^* (q_2, \varepsilon).$$

Každé slovo $w \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$ možno napísať ako $w = uaav$ pre nejaké slová $u, v \in \Sigma^*$. Implikáciu I_2'' dokážeme indukciou vzhľadom na dĺžku slova v.

1. Nech $|v| = 0$. Potom $v = \varepsilon$ a $w = uaa$. Podľa implikácie I_1'' musí platiť $(q_0, ua) \vdash^* (q_1, \varepsilon)$. Keďže ale $q_2 \in \delta(q_1, a)$, dostávame $(q_0, uaa) \vdash^* (q_1, a) \vdash (q_2, \varepsilon)$.
2. Nech tvrdenie platí pre $|v| = k$. Dokážeme, že platí pre $|v| = k + 1$.
   Nech $v \in \Sigma^{k+1}$. Potom $w = uaaxc$, kde $x \in \Sigma^k$ a $c \in \Sigma$. Z indukčného predpokladu potom vyplýva $(q_0, uaax) \vdash^* (q_2, \varepsilon)$. Keďže ale $q_2 \in \delta(q_2, a)$ a $q_2 \in \delta(q_2, b)$, platí $q_2 \in \delta(q_2, c)$, z čoho dostávame $(q_0, uaaxc) \vdash^* (q_2, c) \vdash (q_2, c)$ a implikácia platí.

Ostáva už len využiť platnosť invariantov I_0 až I_2 na dôkaz samotného tvrdenia $L(A) = L$ – ten už je pomerne triviálnou záležitosťou:

$\subseteq$: Nech $w \in L(A)$. Keďže q_2 je jediný akceptačný stav automatu A, musí pre slovo w platiť $(q_0, w) \vdash^* (q_2, \varepsilon)$. Podľa invariantu I_2 potom $w \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$, a teda $w \in L$.

$\supseteq$: Nech $w \in L$. To možno napísať aj ako $w \in \Sigma^* \cdot \{aa\} \cdot \Sigma^*$ a z invariantu I_2 potom vyplýva $(q_0, w) \vdash^* (q_2, \varepsilon)$, a teda $w \in L(A)$.

Tvrdenie je dokázané. $\square$

„Odepsilonovanie“ nedeterministických konečných automatov

Na prednáške bolo dokázané, že prechody na ε nie sú pre nedeterministické konečné automaty nevyhnutné: ku každému NKA existuje ekvivalentný, ktorý žiadne nemá.

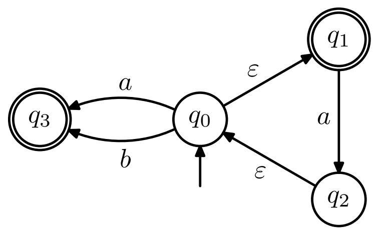
Veta 1. *Nech $A = (K, \Sigma, \delta, q_0, F)$ je nedeterministický konečný automat. Potom existuje nedeterministický konečný automat $A' = (K', \Sigma', \delta', q_0', F')$ taký, že $L(A') = L(A)$ a pre všetky $q \in K'$ je $\delta'(q, \varepsilon) = \emptyset$.*

Konštrukcií zbavujúcich nedeterministický konečný automat prechodov na prázdne slovo existuje viacero. V nasledujúcom uvedieme algoritmus, ktorého správnosť bola dokázaná na prednáške. Jeho vstupom je nedeterministický konečný automat $A = (K, \Sigma, \delta, q_0, F)$.

1. Pre každý stav $q \in K$ nájdi „epsilonový chvost stavu q“, čo je množina $[q]_\varepsilon$ stavov dosiahnuteľných zo stavu q na prázdne slovo: $[q]_\varepsilon = \{p \in K \mid (q, \varepsilon) \vdash^* (p, \varepsilon)\}$. To možno urobiť napríklad prehľadávaním do hĺbky.
2. Odstráň všetky prechody na ε.
3. Pre každý stav $q \in K$ okrem q_0 a každé $c \in \Sigma$ pridaj prechody zo stavu q na písmeno c do všetkých stavov $p \in K$ takých, že existuje $q' \in \delta(q, c)$ také, že $p \in [q']_\varepsilon$.
4. Pre každé $c \in \Sigma$ pridaj prechody zo stavu q_0 na písmeno c do všetkých stavov $p \in K$ takých, že existujú stavy $q \in [q_0]_\varepsilon$ a $q' \in \delta(q, c)$ také, že $p \in [q']_\varepsilon$.
5. Ak „epsilonový chvost“ $[q_0]_\varepsilon$ obsahuje aspoň jeden akceptačný stav, pridaj q_0 do množiny akceptačných stavov.

Poznámka 2. Treba zdôrazniť, že v bodoch 3 a 4 označuje symbol δ prechodovú funkciu *pôvodného* automatu A. Napríklad v kroku 4 teda netreba uvažovať prechody pridané v kroku 3.

Úloha 3. Uvažujme nedeterministický konečný automat $A = (K, \Sigma, \delta, q_0, F)$ s $K = \{q_0, q_1, q_2, q_3\}$, $\Sigma = \{a, b\}$, $F = \{q_1, q_3\}$ a prechodovou funkciou δ danou prechodovým diagramom na obrázku 3. Štandardnou konštrukciou zbavte automat A prechodov na ε.



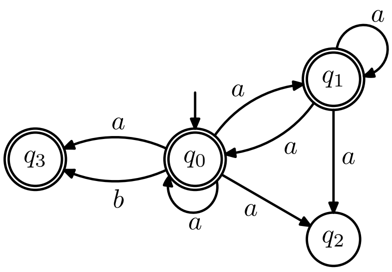
Obr. 3: Prechodový diagram nedeterministického konečného automatu A.

Riešenie. Pre každý stav $q \in K$ najprv nájdeme jeho „epsilonový chvost“ $[q]_\varepsilon$:

$$\begin{aligned}
[q_0]_\varepsilon &= \{q_0, q_1\}, \\
[q_1]_\varepsilon &= \{q_1\}, \\
[q_2]_\varepsilon &= \{q_2, q_0, q_1\}, \\
[q_3]_\varepsilon &= \{q_3\}.
\end{aligned}$$

V kroku 2 odstránime prechody na ε vedúce z q_0 do q_1 a z q_2 do q_0. Pokračujme krokom 3. Keďže $q_2 \in \delta(q_1, a)$ a stavy q_0 a q_1 patria do $[q_2]_\varepsilon$, vo výslednom automate musia byť prechody zo stavu q_1 na písmeno a vedúce do stavov q_0 a q_1. V kroku 3 nepribudnú žiadne ďalšie prechody. Keďže $q_1 \in [q_0]_\varepsilon$, $q_2 \in \delta(q_1, a)$ a stavy q_2, q_0 a q_1 sú všetky v $[q_2]_\varepsilon$, v kroku 4 pribudnú prechody zo stavu q_0 na písmeno a do stavov q_2, q_0 a q_1.

V poslednom kroku sa musí stať akceptačným aj stav q_0, lebo v jeho „epsilonovom chvoste“ je akceptačný stav q_1. Výsledný automat $A' = (K', \Sigma', \delta', q_0', F')$ je potom znázornený prechodovým diagramom na obrázku 4. □



Obr. 4: Prechodový diagram nedeterministického konečného automatu A'.

Poznámka 3. Variantov horeuvedeného algoritmu na „odepsilonovanie“ nedeterministických konečných automatov existuje pomerne veľa. Algoritmus z prednášky je v zásade založený na nahradzovaní postupností prechodov typu „jeden prechod na písmeno a niekoľko prechodov na ε“ jedným prechodom na písmeno. Alternatívne možno uvažovať aj „opačný“ algoritmus, v ktorom sa jedným prechodom na písmeno nahradzujú postupnosti typu „niekoľko prechodov na ε a jeden prechod na písmeno“. Takýto prístup umožňuje spracovať počiatočný stav q_0 konzistentne s ostatnými stavmi – nie je teda nutná analógia kroku 4. Krok 5 by ale naopak bolo nutné vykonať pre všetky stavy (a nielen pre počiatočný stav q_0).

Obidva spomenuté varianty „odepsilonovacieho“ algoritmu možno upraviť aj pre nedeterministické konečné automaty s viac ako jedným počiatočným stavom, kde ich možno interpretovať azda najprirodzenejším a najelegantnejším spôsobom; podobné úvahy ale prekračujú rámec tohto predmetu.