Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

# Prídavná informácia a zložitosť nedeterministických konečných automatov

Diplomová práca

2017
Bc. Šimon Sádovský

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

# Prídavná informácia a zložitosť nedeterministických konečných automatov

Diplomová práca

Študijný program: Informatika
Študijný odbor: 2508 Informatika
Školiace pracovisko: Katedra informatiky
Školiteľ: prof. RNDr. Branislav Rovan, PhD.

Bratislava, 2017
Bc. Šimon Sádovský

35818858

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

# ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Bc. Šimon Sádovský
**Študijný program:** informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
**Študijný odbor:** informatika
**Typ záverečnej práce:** diplomová
**Jazyk záverečnej práce:** slovenský
**Sekundárny jazyk:** anglický

**Názov:** Prídavná informácia a zložitosť nedeterministických konečných automatov
*Supplementary Information and Complexity of Nondeterministic Finite Automata*

**Cieľ:** Preskúmať užitočnosť prídavnej informácie o vstupnom slove pre zníženie zložiosti nedeterminstických konečných automatov pre akceptáciu jazykov. Práca nadväzuje napredchádzajúce diplomové práce, v ktorých sa skúmal tento problém pre deterministické automaty.

**Vedúci:** prof. RNDr. Branislav Rovan, PhD.
**Katedra:** FMFI.KI - Katedra informatiky
**Vedúci katedry:** prof. RNDr. Martin Škoviera, PhD.

**Spôsob sprístupnenia elektronickej verzie práce:**
bez obmedzenia

**Dátum zadania:** 16.12.2015

**Dátum schválenia:** 16.12.2015

prof. RNDr. Rastislav Kráľovič, PhD.
garant študijného programu

študent

vedúci práce

**Poďakovanie:**

# Abstrakt

V práci skúmame vplyv prídavnej informácie na zložitosť riešenia problému. Ako výpočtový model sme zvolili nedeterministické konečné automaty a mierou zložitosti je počet stavov. Formalizáciou nášho problému je rozklad nedeterministického konečného automatu na dvojicu nedeterministických konečných automatov takých, že jazyk pôvodného automatu je prienikom jazykov týchto dvoch automatov. Navyše očakávame, že oba tieto automaty budú jednoduchšie ako pôvodný automat. V práci dokazujeme rozložiteľnosť respektíve nerozložiteľnosť konkrétnych regulárnych jazykov. Dokazujeme uzáverové a iné vlastnosti tried nedeterministicky rozložiteľných a nedeterministicky nerozložiteľných regulárnych jazykov. Charakterizujeme triedu jazykov tvorených jedným slovom vzhľadom na rozložiteľnosť. Skúmame jazyky, ktorých minimálny nedeterministický automat je tvorený práve jedným cyklom. Ukazujeme rozdiel medzi nedeterministickou a deterministickou rozložiteľnosťou regulárnych jazykov.

**Kľúčové slová:** nedeterministický konečný automat, rozklad nedeterministického konečného automatu, nedeterministická rozložiteľnosť, prídavná informácia, popisná zložitosť

# Abstract

We study the effect of supplementary information on the complexity of problem solution. We have chosen nondeterministic finite automaton as the computational model and we measure the complexity by the number of states. We formalize our problem via decomposition of nondeterministic finite automaton into two nondeterministic finite automata, such that the language accepted by the original automaton is the intersection of languages accepted by this two automata. Moreover, we require both automata in the decomposition to be simpler than the original automaton. We show decomposability and nondecomposability of particular regular languages. We show closure and other properties of classes of nondeterministically decomposable and nondecomposable regular languages. We characterize the class of languages consisting of exactly one word with respect to decomposability. We examine languages accepted by nondeterministic automata consisting of exactly one cycle. We show the difference between nondeterministic and deterministic decomposability of regular languages.

**Keywords:** nondeterministic finite automaton, decomposition of nondeterministic finite automaton, nondeterministic decomposability, supplementary information, descriptional complexity

# Obsah

# Zoznam obrázkov

# Úvod

Konečný automat je jednoduchý výpočtový model, ktorý má široké praktické uplatnenie. Pojem konečného automatu prvý krát zaviedli McCulloch a Pitts v [McCulloch and Pitts, 1943]. Odvtedy bolo študovaných mnoho formalizácií tohto pojmu, ktoré môžeme rozdeliť do dvoch základných skupín: prekladače a akceptory. Ústredným pojmom našej práce je nedeterministický konečný automat, ktorý patrí medzi akceptory.

Našou motiváciou je otázka užitočnosti prídavnej informácie pri akceptovaní jazyka. Voľne povedané, ak automatu našepkám, že vstup, ktorý ide rozpoznávať, patrí do nejakého poradného jazyka, viem tým dosiahnuť, že na rozpoznávanie pôvodného jazyka stačí automat menšej zložitosti? Uveďme jeden príklad. Uvažujme, že chceme rozpoznávať jazyk $\{w \in \{a\}^* \mid |w| \equiv 0 \ (mod\ 6)\}$ a chceme ho rozpoznávať nedeterministickým konečným automatom. Ľahko vidno, že minimálny nedeterministický konečný automat pre tento jazyk má 6 stavov. Čo ak automatu našepkám, že dĺžka vstupu je deliteľná tromi? Vtedy nám stačí použiť automat s dvomi stavmi.

Iný pohľad na formalizáciu tejto otázky je, či vieme rozložiť automat rozpoznávajúci jazyk na dva, ktoré sú nejakým spôsobom jednoduchšie ako pôvodný automat, pričom prienik jazykov, ktoré rozpoznávajú jednotlivé jednoduchšie automaty je pôvodný jazyk. Na takýto rozklad sa môžeme pozerať tak, že jeden z automatov v rozklade poskytuje prídavnú informáciu a druhý je zjednodušením pôvodného automatu. Teda jazyk rozpoznávaný jedným z týchto dvoch automatov plní funkciu poradného jazyka.

Spomeňme ešte, že otázka užitočnosti prídavnej informácie sa dá takto definovať pre akýkoľvek výpočtový model, nie nutne iba pre konečné automaty. V našej práci budeme tento problém skúmať výlučne pre nedeterministické konečné automaty. V minulosti bol tento problém už skúmaný na našej fakulte pre deterministické konečné automaty v práci [Gaži and Rovan, 2008] a pre deterministické zásobníkové automaty v práci [Labath and Rovan, 2011].

V Kapitole 1 definujeme potrebné pojmy, ktoré potrebujeme v našej práci. Takisto uvádzame potrebné výsledky prevzaté z literatúry.

V Kapitole 2 skúmame konkrétne jazyky vzhľadom na rozložiteľnosť, budujeme dôkazové techniky a repertoár tvrdení potrebných v ďalšom texte.

Zaujímavou otázkou je, či je pojem rozložiteľnosti rôzny ak uvažujeme deterministické respektíve nedeterministické konečné automaty. Túto otázku riešime v Kapitole

3. Uvádzame nekonečnú postupnosť jazykov takú, že každý z týchto jazykov je nedeterministicky nerozložiteľný a deterministicky rozložiteľný.

V Kapitole 4 skúmame uzáverové vlastnosti tried rozložiteľných a nerozložiteľných jazykov. Ukazujeme, že príliš malé nedeterministické konečné automaty sú nerozložiteľné. Ukazujeme, že vloženie nového symbolu do jazyka nezmení rozložiteľnosť resp. nerozložiteľnosť jazyka. Charakterizujeme jazyky tvorené jedným slovom vzhľadom na rozložiteľnosť. Skúmame jazyky, ktorých minimálny nedeterministický konečný automat je tvorený práve jedným cyklom.

# Kapitola 1

# Definície, potrebné výsledky

V kapitole definujeme pojmy, zavádzame označenia a uvádzame výsledky potrebné pre našu prácu.

## 1.1 Nedeterministický konečný automat

Nedeterministický konečný automat je dobre známy model, avšak existuje viac jeho ekvivalentných definícii, preto uvádzame tú, ktorú budeme používať v našom texte.

**Definícia 1.1.1.** ***Nedeterministický konečný automat*** *je pätica* $(K, \Sigma, \delta, q_0, F)$, *kde:*

1. *$K$ je konečná množina stavov*
2. *$\Sigma$ je konečná vstupná abeceda*
3. *$q_0 \in K$ je počiatočný stav*
4. *$F \subseteq K$ je množina akceptačných stavov*
5. *$\delta : K \times (\Sigma \cup \{\varepsilon\}) \to 2^K$ je prechodová funkcia*

**Poznámka 1.1.1.** *Nedeterministický konečný automat sa skrátene označuje NKA.*

**Poznámka 1.1.2.** *Ak v texte hovoríme o nejakom automate $A$, štandardne berieme, že $A = (K_A, \Sigma_A, \delta_A, q_{0A}, F_A)$ a teda ak hovoríme o množine $K_A$, myslíme tým množinu stavov automatu $A$. Analogicky to platí aj pre $\Sigma_A, \delta_A, q_{0A}, F_A$. Pokial je z kontextu jasné, o ktorý automat sa jedná, dolný index $A$ vynechávame a píšeme skrátene $K, \Sigma, \delta, q_0, F$.*

**Definícia 1.1.2.** ***Konfigurácia*** *nedeterministického konečného automatu $A$ je dvojica $(q, u) \in K \times \Sigma^*$, kde $q$ je stav, v ktorom sa automat nachádza a $u$ je ešte nedočítaná časť vstupného slova.*

**Definícia 1.1.3.** ***Krok výpočtu*** *nedeterministického konečného automatu A je relácia* $\vdash_A$ *na konfiguráciách definovaná* $(q, au) \vdash_A (p, u) \Leftrightarrow p \in \delta(q, a)$, $q, p \in K, u \in \Sigma^*, a \in \Sigma \cup \{\varepsilon\}$. *Reflexívno-tranzitívny uzáver relácie* $\vdash_A$ *označujeme* $\vdash_A^*$. *Ak je z kontextu jasné, o ktorý konečný automat sa jedná, index A vynechávame a píšeme iba* $\vdash$.

**Definícia 1.1.4.** ***Jazyk*** *akceptovaný (definovaný) nedeterministickým konečným automatom A je jazyk* $L(A) = \{w \in \Sigma^* | \exists q_F \in F : (q_0, w) \vdash^* (q_F, \varepsilon)\}$.

**Definícia 1.1.5.** ***Stavovou zložitosťou*** *nedeterministického konečného automatu A (označujeme* $\#_S(A)$*) rozumieme počet jeho stavov, t.j.* $\#_S(A) = |K|$.

**Definícia 1.1.6.** ***Nedeterministickú stavovú zložitosť*** *jazyka* $L \in \mathscr{R}$ *(označujeme nsc(L) - z anglického nondeterministic state complexity) definujeme* $nsc(L) = min\{\#_S(A) | L(A) = L\}$.

**Definícia 1.1.7.** *Nech* $L \in \mathscr{R}$. ***Minimálnym nedeterministickým konečným automatom pre jazyk L*** *rozumieme ľubovolný nedeterministický konečný automat A taký, že* $\#_S(A) = nsc(L)$.

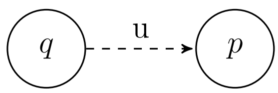
## 1.2 Ďalšie označenia

**Označenie 1.2.1.** ***Dĺžku slova*** *w označujeme* $|w|$.

V práci často uvádzame NKA pomocou štandardného stavového diagramu. Pre lepšiu čitateľnosť dôkazov zavádzame nasledovné označenia.

**Označenie 1.2.2.** *Nech u je ľubovolné slovo,* $k \in \mathbb{N}$. *Potom* $pref(u, k)$ *označujeme* ***prefix slova*** *u* ***dĺžky*** *k a* $suff(u, k)$ *označujeme* ***suffix slova*** *u* ***dĺžky*** *k.*

**Označenie 1.2.3.** *Nech* $u = u_1 u_2 \dots u_n$ *je ľubovolné slovo. Ak v diagrame NKA A použijeme označenie prechodu slovom:*




*myslíme tým, že v automate A sa dá zo stavu q dostať do stavu p na slovo u pričom zo stavov, v ktorých sa automat A nachádza počas čítania slova u sa nedá už nikam inam dostať. Formálne, existujú* $q_0, q_1, \dots, q_n \in K_A$ *také, že* $q_0 = q, q_n = p, \delta_A(q, u_1) \ni q_1$ *a pre* $0 < i < n$ *platí* $\delta_A(q_i, u_{i+1}) = \{q_{i+1}\}, q_i \notin F_A$.

## 1.3 Definícia problému

Na základe úvah uvedených v Úvode našej práce zavedieme ústredné pojmy našej práce.

**Definícia 1.3.1.** *Nech $A$ je nedeterministický konečný automat. Potom dva nedeterministické konečné automaty $A_1, A_2$ také, že $L(A) = L(A_1) \cap L(A_2)$ nazveme* ***rozklad automatu*** *$A$. Ak navyše platí $\#_S(A_1) < \#_S(A)$ a $\#_S(A_2) < \#_S(A)$, nazývame tento rozklad* ***netriviálny***. *Ak existuje netriviálny rozklad automatu $A$, tak automat $A$ nazývame* ***rozložiteľný***.

**Definícia 1.3.2.** *Nech $L \in \mathscr{R}$ a $A$ je nejaký minimálny NKA pre jazyk $L$.* ***Jazyk*** *$L$ nazývame* ***nedeterministicky rozložiteľný*** *práve vtedy, keď je automat $A$ rozložiteľný.*

Ukážeme, že vlastnosť jazyka byť rozložiteľný je dobre definovaná, teda nezávisí od výberu minimálneho konečného automatu pre jazyk.

**Tvrdenie 1.3.1.** *Nech $A_1$ a $A_2$ sú minimálne NKA pre jazyk $L$. Potom $A_1$ je rozložiteľný práve vtedy keď je $A_2$ rozložiteľný.*

*Dôkaz.* Uvažujme ľubovolný jazyk $L \in \mathscr{R}$. Ak existuje pre daný jazyk unikátny minimálny NKA, tak niet čo dokazovať. Uvažujme teda, že pre jazyk $L$ existuje viacero minimálnych NKA. Nech $A_1$ a $A_2$ sú rôzne minimálne NKA pre jazyk $L$. Dokážeme, že automat $A_1$ je rozložiteľný práve vtedy, keď je rozložiteľný automat $A_2$. Nech teda existuje netriviálny rozklad automatu $A_1$. Teda existujú NKA $B_1$ a $B_2$ také, že $L(B_1) \cap L(B_2) = L(A_1) = L$ a $\#_S(B_1) < \#_S(A_1)$, $\#_S(B_2) < \#_S(A_1)$. Nakoľko $A_1$ a $A_2$ sú oba minimálne automaty pre jazyk $L$, tak platí $\#_S(A_1) = \#_S(A_2)$ a $L(A_1) = L(A_2) = L$. Teda platí $\#_S(B_1) < \#_S(A_2)$, $\#_S(B_2) < \#_S(A_2)$ a taktiež $L(B_1) \cap L(B_2) = L(A_2) = L$, teda $B_1$ a $B_2$ tvoria zároveň netriviálny rozklad automatu $A_2$. Daná úvaha sa dá analogicky spraviť aj opačným smerom a dokázať, že ak je rozložiteľný automat $A_2$, tak potom je rozložiteľný aj automat $A_1$. □

**Poznámka 1.3.1.** *V našej práci budeme takmer vždy hovoriť o nedeterministickej rozložiteľnosti jazyka, preto budeme písať skrátene o rozložiteľnosti jazyka. Plný výraz nedeterministická rozložiteľnosť jazyka budeme používať iba v prípadoch, keď bude treba zvýrazniť, že ide práve o nedeterministickú rozložiteľnosť a nie deterministickú.*

Ľahko vidno, že rozklad NKA $A$ existuje vždy a tvorí ho samotný automat $A$ a NKA pre jazyk $\Sigma_A^*$. Samozrejme tento rozklad nie je netriviálny a rovnako nie je ani ničím zaujímavý. Preto nás bude v prípade automatov zaujímať, za akých podmienok existuje ich netriviálny rozklad. Pri jazykoch nás bude zaujímať, či sú rozložiteľné.

Zmysel nasledujúcej lemy je v zjednodušení dôkazov niektorých tvrdení v našej práci, kde potrebujeme predpokladať existenciu rozkladu netriviálneho rozkladu nejakého automatu a následne dokázať niečo o výpočtoch NKA ktoré tvoria tento rozklad.

Vďaka tejto Leme môžeme predpokladať, že dané výpočty v každom kroku spracujú nejaký znak zo vstupu, čo robí dôkazy prehľadnejšími.

**Lema 1.3.1** (o bezepsilonových NKA)**.** *Nech $A$ je NKA. Potom platia nasledovné tvrdenia.*

*(a) existuje NKA $A'$ taký, že $L(A') = L(A)$, $\#_S(A) = \#_S(A')$ a automat $A'$ neobsahuje prechody na $\varepsilon$*

*(b) ak je $A$ rozložitelný, potom existuje netriviálny rozklad automatu $A$ na NKA $A_1^\varepsilon, A_2^\varepsilon$ taký, že $A_1^\varepsilon$ a $A_2^\varepsilon$ neobsahujú prechody na $\varepsilon$*

*Dôkaz.* Tvrdenie (a) vyplýva priamo zo štandardnej konštrukcie odepsilonovaného NKA k ľubovoľnému NKA.

Dokážeme tvrdenie (b). Automat $A$ je rozložiteľný, to znamená, že existuje netriviálny rozklad automatu $A$ na automaty $A_1$ a $A_2$, čo znamená, že $L(A) = L(A_1) \cap L(A_2), \#_S(A_1) < \#_S(A), \#_S(A_2) < \#_S(A)$. Podľa (a) však existujú automaty $A_1'$ a $A_2'$ také, že $L(A_1') = L(A_1)$, $\#_S(A_1) = \#_S(A_1')$ a $L(A_2') = L(A_2)$, $\#_S(A_2) = \#_S(A_2')$ pričom navyše automaty $A_1'$ a $A_2'$ neobsahujú prechody na $\varepsilon$. To však znamená, že $L(A) = L(A_1') \cap L(A_2'), \#_S(A_1') < \#_S(A), \#_S(A_2') < \#_S(A)$, teda $A_1'$ a $A_2'$ tvoria taktiež netriviálny rozklad automatu $A$. Teda stačí položiť $A_1^\varepsilon = A_1', A_2^\varepsilon = A_2'$. □

## 1.4 Techniky určovania dolnej hranice počtu stavov NKA

Na skúmanie otázky rozložitelnosti jazyka musíme mať nástroje, pomocou ktorých vieme k jazykom hladať ich minimálne automaty. V nasledujúcej časti uvedieme techniky, pomocou ktorých budeme schopní určovať dolné hranice pre počet stavov nedeterministického konečného automatu pre daný jazyk. Pre deterministické konečné automaty máme k dispozícii `Myhill-Nerodovú vetu`, ktorá vždy dokáže určiť tesnú spodnú hranicu pre počet stavov potrebných pre deterministický konečný automat rozpoznávajúci daný jazyk. Pri nedeterministických konečných automatoch je situácia horšia. Takúto silnú techniku nemáme k dispozícii. Avšak máme k dispozícii techniky, ktoré nám poskytujú aspoň nejaké, nie nutne tesné, dolné hranice pre počet stavov potrebných pre nedeterministický konečný automat rozpoznávajúci daný jazyk. Uvádzame dve techniky - Techniku mätúcich množín (z anglického Fooling set technique) a techniku rozšírených mätúcich množín (z anglického Extended fooling set technique), ktoré čerpáme z [Palioudakis, 2012] a [Glaister and Shallit, 1996].

**Definícia 1.4.1** (Mätúca množina)**.** *Nech $L$ je jazyk, $n \in \mathbb{N}$. Nech $P = \{(x_i, y_i) | 1 \leq i \leq n\}$ taká, že:*

*(a)* $x_i y_i \in L$ *pre* $1 \leq i \leq n$

*(b)* $x_i y_j \notin L$ *pre* $1 \leq i, j \leq n$ *a* $i \neq j$

*Potom množinu* $P$ *nazývame* ***mätúca množina pre jazyk*** $L$.

**Veta 1.4.1** (Technika mätúcich množín)**.** *Nech* $L$ *je regulárny jazyk a existuje mätúca množina* $P$ *pre jazyk* $L$*. Potom každý NKA akceptujúci* $L$ *má aspoň* $|P|$ *stavov (t.j.* $nsc(L) \geq |P|$*).*

*Dôkaz.* Aby sme nahliadli, čo je za touto technikou, uvedieme aj dôkaz. Označme $|P| = n$ a postupujme sporom. Nech platia predpoklady tvrdenia a nech existuje NKA A ktorý má menej stavov ako $n$. Pozrime sa na výpočty automatu $A$ na slovách $x_i y_i$ pre $1 \leq i \leq n$. Podľa definície množiny $P$ musí platiť $(q_{0_A}, x_i y_i) \vdash^* (p_i, y_i) \vdash^* (q_{i_F}, \varepsilon)$ kde $p_i \in K_A$ a $q_{i_F} \in F_A$. Pozrime sa teraz pozornejšie na stavy $p_i$. Nakolko platí, že automat A má menej stavov ako je $n$, musí platiť, že existujú také $k \neq l$, že $p_k = p_l$. Potom však platí, že $(q_{0_A}, x_k y_l) \vdash^* (p_l, y_l) \vdash^* (q_{i_F}, \varepsilon)$. Potom však $x_k y_l \in L$ čo je spor s definíciou množiny P. Teda A má aspoň n stavov. □

Drobnou úpravou tejto vety dostaneme silnejšie tvrdenie.

**Definícia 1.4.2** (Rozšírená mätúca množina)**.** *Nech* $L$ *je jazyk. Nech* $n \in \mathbb{N}$*. Nech* $P = \{(x_i, y_i) | 1 \leq i \leq n\}$ *taká, že:*

*(a)* $x_i y_i \in L$ *pre* $1 \leq i \leq n$

*(b)* $x_i y_j \notin L$ *alebo* $x_j y_i \notin L$ *pre* $1 \leq i, j \leq n$ *a* $i \neq j$

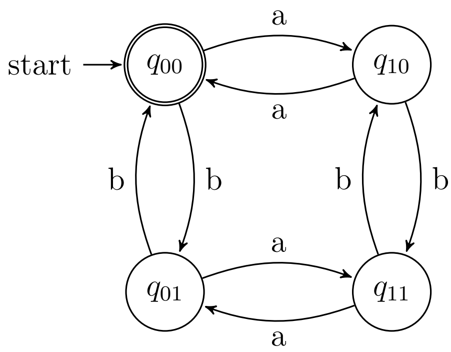
*Potom množinu* $P$ *nazývame* ***rozšírená mätúca množina pre jazyk*** $L$.

**Veta 1.4.2** (Technika rozšírených mätúcich množín)**.** *Nech* $L$ *je regulárny jazyk a existuje rozšírená mätúca množina* $P$ *pre jazyk* $L$*. Potom každý NKA akceptujúci* $L$ *má aspoň* $|P|$ *stavov (t.j.* $nsc(L) \geq |P|$*).*

Dôkaz je takmer identický ako dôkaz pre 1.4.1 a je triviálne ho rozšíriť tak, aby dokazoval toto tvrdenie, preto ho neuvádzame. Takisto je ľahko vidno, že ak je množina mätúcou množinou pre jazyk $L$, je aj rozšírenou mätúcou množinou pre $L$.

Prirodzená otázka, ktorá sa ponúka, je: „Ako nájsť čo najväčšiu (rozšírenú) mätúcu množinu pre daný jazyk $L$? “. Algoritmus, pomocou ktorého by sa táto množina dala skonštruovať známy nie je, avšak v [Glaister and Shallit, 1996] autori ponúkajú nasledujúcu heuristiku, ktorá, ako sa zdá, často zafunguje velmi dobre. Najprv skonštruujme NKA akceptujúci jazyk $L$. Nech pre každý stav $q$ tohto automatu je $x_q$ najkratšie slovo také, že platí $(q_0, x_q) \vdash^* (q, \varepsilon)$ a nech $y_q$ je najkratšie slovo také, že platí $(q, y_q) \vdash^* (q_F, \varepsilon)$, kde $q_F$ je akceptačný stav. Potom zvoľme P ako nejakú vhodnú podmnožinu $\{(x_q, y_q) | q \in K\}$.

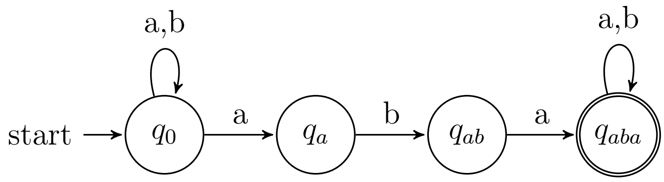
**Príklad 1.4.1.** Uvažujme jazyk $L = \{w \in \{a, b\}^* \mid \#_a(w) \equiv 0\ (mod\ 2) \wedge \#_b(w) \equiv 0\ (mod\ 2)\}$. NKA akceptujúci jazyk $L$ uvádzame pomocou diagramu.




Obr. 1.1: NKA akceptujúci jazyk $L$

Teraz, použijúc techniky uvedené v predošlom, dokážeme, že tento NKA je minimálnym NKA pre jazyk $L$. Uvažujme množinu dvojíc slov $F = \{(\varepsilon, \varepsilon), (a, a), (ab, ab), (b, b)\}$. Množina $F$ je podla definície 1.4.1 mätúcou množinou pre jazyk $L$. Nakoľko $|F| = 4$, tak podla vety 1.4.1 platí $nsc(L) \geq 4$. Keďže sa nám podarilo zostrojiť NKA akceptujúci $L$, ktorý má práve 4 stavy, tak tento NKA je minimálnym automatom pre jazyk $L$, t.j. $nsc(L) = 4$.

**Príklad 1.4.2.** Uvažujme jazyk $L = \{w_1abaw_2 \mid w_1, w_2 \in \{a, b\}^*\}$. NKA akceptujúci jazyk $L$ uvádzame pomocou diagramu.




Obr. 1.2: NKA akceptujúci jazyk $L$

Použijúc techniky uvedené v predošlom dokážeme, že tento NKA je minimálny NKA pre jazyk $L$. Uvažujme množinu dvojíc slov $F = \{(\varepsilon, aba), (a, ba), (ab, a), (aba, \varepsilon)\}$. Množina $F$ je podla definície 1.4.2 rozšírenou mätúcou množinou pre jazyk $L$. Nakoľko $|F| = 4$, tak podla vety 1.4.2 platí $nsc(L) \geq 4$. Keďže sa nám podarilo zostrojiť NKA akceptujúci $L$, ktorý má práve 4 stavy, tak tento NKA je minimálny NKA pre jazyk $L$, t.j. $nsc(L) = 4$. Ešte spomeňme, že pri dokazovaní minimality pomocou techniky mätúcich množín (nie rozšírených) by sme neuspeli, nakoľko najväčšia možná mätúca množina pre jazyk $L$ obsahuje 2 prvky.

# Kapitola 2

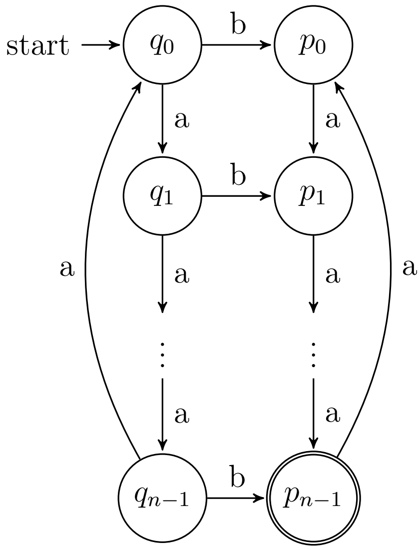
# Rozložitelné a nerozložitelné jazyky

V kapitole skúmame konkrétne jazyk vzhľadom na ich rozložiteľnosť. Cieľom kapitoly je poskytnúť základný vhlad do problematiky a takisto vybudovať repertoár tvrdení, ktoré budeme používať v ďalšom texte pri dôkazoch.

## 2.1 Rozložitelné jazyky

**Tvrdenie 2.1.1.** *Uvažujme jazyky* $L_n = \{a^k ba^l | (l + k) \equiv 0 (mod\ n)\}$. *Ak* $n \geq 2$, *potom jazyk* $L_n$ *je rozložiteľný.*

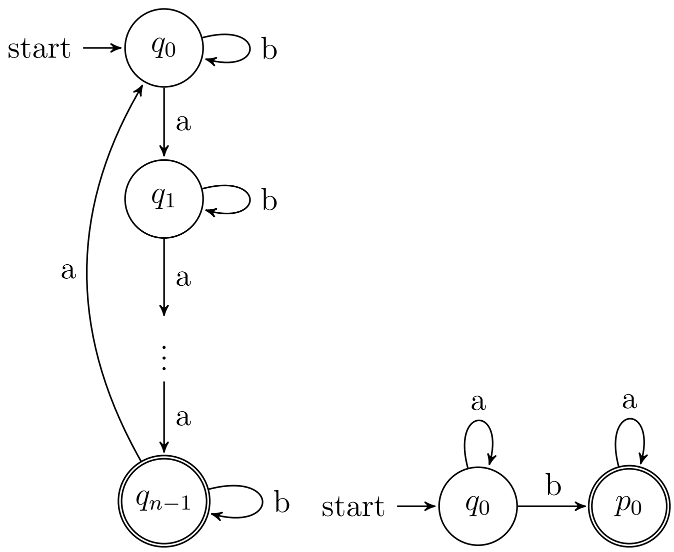
*Dôkaz.* Uvažujme $n \in \mathbb{N}, n \geq 2$. Aby sme dokázali, že jazyk je regulárny a teda má význam uvažovať o jeho rozklade, zostrojme NKA $A_n$ taký, že $L(A_n) = L_n$. Hľadaný NKA uvádzame pomocou diagramu.




Obr. 2.1: automat $A_n$ pre jazyk $\{a^k ba^l | (l + k) \equiv 0 (mod\ n)\}$

Uvažujme množinu dvojíc slov $F_n = \{(a^l, ba^{n-l}), (a^l b, a^{n-l}) \mid 0 \leq l \leq n - 1\}$. Podla definície 1.4.2 je množina $F_n$ rozšírenou mätúcou množinou pre jazyk $L_n$. $|F_n| = 2n$, teda podľa Vety 1.4.2 $nsc(L_n) \geq 2n$. Keďže $L(A_n) = L_n$ a $\#_S(A_n) = n + 2$, tak $nsc(L_n) = 2n$ a automat $A_n$ je minimálny NKA pre jazyk $L_n$.

Teraz zostrojme netriviálny rozklad automatu $A_n$. Hľadané NKA $A_n^1$ a $A_n^2$ uvádzame pomocou ich diagramov.

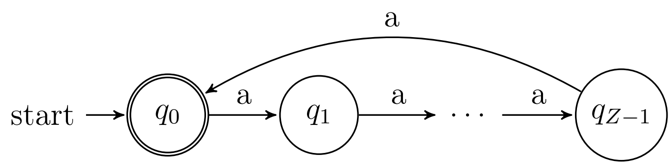



Obr. 2.2: rozklad automatu $A_n$

Ľahko vidno, že uvedené NKA pre $n \geq 2$ tvoria netriviálny rozklad automatu $A_n$, teda že platí $\#_S(A_n^1) < 2n$, $\#_S(A_n^2) < 2n$, $L(A_n^1) \cap L(A_n^2) = L(A_n)$. □

**Veta 2.1.1.** *Uvažujme jazyky $L_Z = \{a^{kZ} \mid k \in \mathbb{N}\}$ pre $Z \in \mathbb{N}$. Ak $Z$ nie je mocninou prvočísla, potom jazyk $L_Z$ je rozložiteľný.*

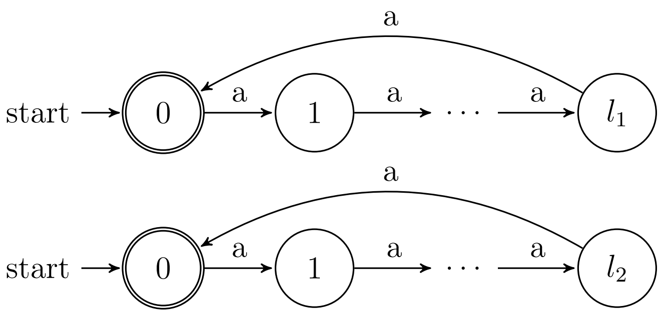
*Dôkaz.* Podľa predpokladu vety uvažujme $Z \in \mathbb{N}, Z > 0$, $Z$ nie je mocninou prvočísla. Najprv ukážeme, že $nsc(L_Z) = Z$. Zostrojme NKA $A_Z$ taký, že $L(A_Z) = L_Z$. Automat uvádzame pomocou diagramu.




Obr. 2.3: automat $A_Z$

Uvažujme množinu dvojíc slov $F_Z = \{(a^i, a^{Z-i}) \mid 0 \leq i \leq Z - 1\}$. Podla definície 1.4.1 je množina $F_Z$ mätúcou množinou pre jazyk $L_Z$. Nakoľko $|F_Z| = Z$, tak podľa Vety 1.4.1 $nsc(L_Z) \geq Z$. Nakoľko $L(A_Z) = L_Z$ a $\#_S(A_Z) = Z$, tak platí $nsc(L_Z) = Z$. Intuitívne je jasné, že automat „počíta zvyšok po delení $Z$“.

Teraz nájdeme netriviálny rozklad automatu $A_Z$. Nech $p_1^{m_1} p_2^{m_2} ... p_r^{m_r}$ je prvočíselný rozklad čísla $Z$. Podla predpokladov vety platí, že $r \geq 2$. Najprv načrtneme intuitívny pohľad vyplývajúci z vlastností zložených čísel a potom túto intuíciu sformalizujeme. Automaty v rozklade budú počítať zvyšok po delení $p_1^{m_1}$ a zvyšok po delení $p_2^{m_2}...p_r^{m_r}$ a budú akceptovať, ak nimi počítaný zvyšok vyjde 0. Ak oba zvyšky vyjdú 0, tak dostaneme slovo, v ktorom počet písmen $a$ je deliteľný $p_1^{m_1}$ a zároveň je deliteľný $p_2^{m_2}...p_r^{m_r}$. Nakoľko $p_1, p_2, ..., p_r$ sú navzájom rôzne prvočísla, tak potom počet písmen $a$ v zmienenom slove je deliteľný $Z = p_1^{m_1}p_2^{m_2}...p_r^{m_r}$. Teraz uveďme hľadané automaty, ktoré tvoria rozklad automatu $A_Z$. Automaty uvádzame pomocou diagramov. Pre prehľadnosť diagramov zaveďme označenie $l_1 = p_1^{m_1}$ a $l_2 = p_2^{m_2}...p_r^{m_r}$




Obr. 2.4: rozklad automatu $A_Z$ na automaty $A_1^Z$(hore) a $A_2^Z$(dole)

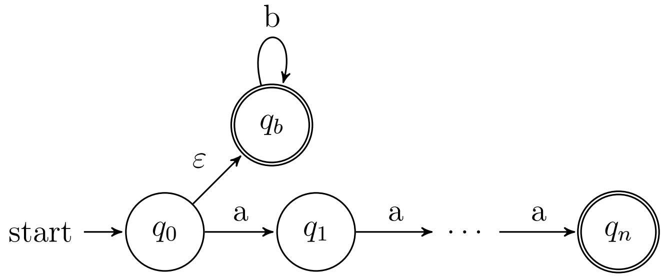
Automaty v rozklade označme $A_1^Z$ a $A_2^Z$. Formálne dokážme, že $L(A_1^Z) \cap L(A_2^Z) = L(A_Z)$.
$\subseteq$ : Nech $w \in L(A_1^Z) \cap L(A_2^Z)$. Z konštrukcie automatov $A_1$ a $A_2$ vyplýva, že slovo $w$ obsahuje iba znaky $a$ a jeho dĺžka je deliteľná $p_1^{m_1}$ a zároveň je deliteľná $p_2^{m_2}...p_r^{m_r}$. Z toho vyplýva, že $\exists t \in \mathbb{N} : w = a^{tp_1^{m_1}p_2^{m_2}...p_r^{m_r}}$. A teda $w \in L(A_Z)$.
$\supseteq$ : Nech $w \in L(A_Z)$. Teda $\exists t \in \mathbb{N} : w = a^{tp_1^{m_1}p_2^{m_2}...p_r^{m_r}}$. Nakoľko $L(A_1^Z) = \{a^{kp_1^{m_1}} | k \in \mathbb{N}\}$, tak $w \in L(A_1^Z)$. Nakoľko $L(A_2^Z) = \{a^{kp_2^{m_2}...p_r^{m_r}} | k \in \mathbb{N}\}$, tak $w \in L(A_2^Z)$. Z toho $w \in L(A_1^Z) \cap L(A_2^Z)$.

Nakoľko $\#_S(A_1^Z) < \#_S(A_Z)$ a $\#_S(A_2^Z) < \#_S(A_Z)$, tento rozklad je netriviálny, čím je tvrdenie dokázané.

□

**Tvrdenie 2.1.2.** *Uvažujme jazyky $L_n = \{a^n\} \cup \{b\}^*$. Ak $n \geq 2$, potom jazyk $L_n$ je rozložiteľný.*

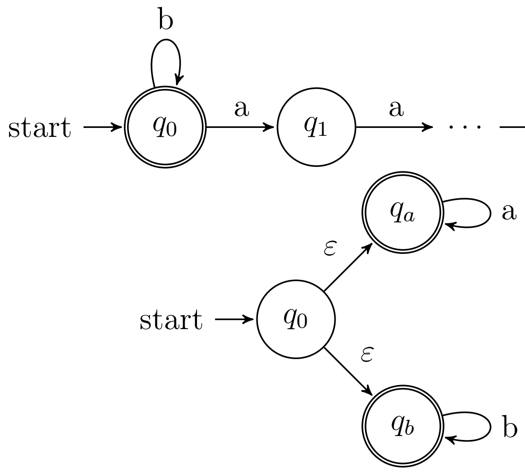
*Dôkaz.* Podľa prepokladu uvažujme $n \geq 2$. Najprv dokážeme, že $nsc(L_n) = n + 2$. Najprv zostrojme NKA $A_n$ akceptujúci jazyk $L_n$. Automat $A_n$ uvádzame pomocou diagramu.




Obr. 2.5: automat $A_n$ pre jazyk $\{a^n\} \cup \{b\}^*$

Uvažujme množinu dvojíc slov $F_n = \{(b, b)\} \cup \{(a^i, a^{n-1}) | 0 \leq i \leq n\}$. Táto množina je podľa definície 1.4.2 rozšírenou mätúcou množinou pre jazyk $L_n$. Keďže $|F_n| = n+2$, tak podľa Vety 1.4.2 $nsc(L_n) \geq n+2$. Nakoľko automat $L(A_n)$ a $\#_S(A_n) = n+2$, tak $nsc(L_n) = n+2$ a automat $A_n$ je minimálny NKA pre jazyk $L_n$.

Teraz zostrojíme netriviálny rozklad automatu $A_n$, čím skompletizujeme dôkaz. Rozklad uvádzame pomocou diagramu.

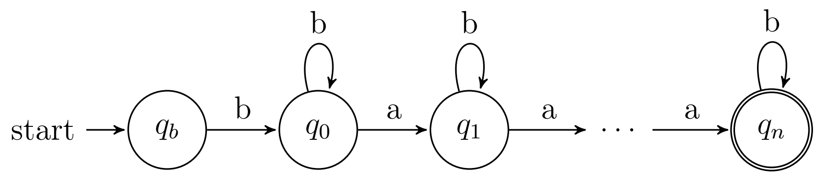



Obr. 2.6: netriviálny rozklad automatu $A_n$ z Obr. 2.5 na automaty $A_1^n$(hore) a $A_2^n$(dole)

$L(A_1^n) = \{b^k, b^k a^n | k \in \mathbb{N}\}$, $L(A_2^n) = \{a\}^* \cup \{b\}^*$. Teda $L(A_1^n) \cap L(A_2^n) = L(A_n)$. Nakoľko $\#_S(A_1^n) < \#_S(A_n)$ a $\#_S(A_2^n) < \#_S(A_n)$, automaty $A_1^n$ a $A_2^n$ tvoria netriviálny rozklad automatu $A_n$. $\square$

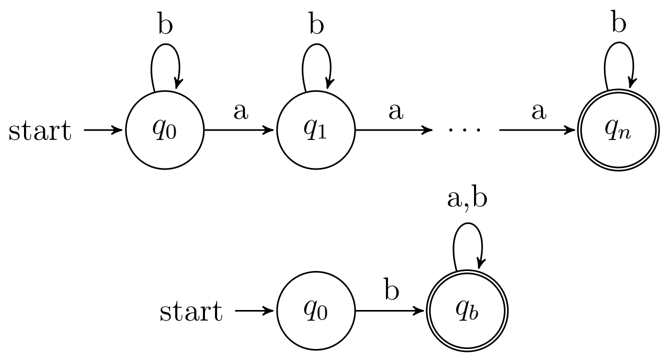
**Tvrdenie 2.1.3.** *Uvažujme jazyky* $L_n = \{b\}.\{w \in \{a,b\}^* | \#_a(w) = n\}$. *Ak* $n \geq 1$ , *potom je jazyk* $L_n$ *rozložiteľný.*

*Dôkaz.* Uvažujme $n \in \mathbb{N}, n \geq 1$. Ukážeme, že $nsc(L_n) = n+2$. Najprv zostrojíme NKA $A_n$ pre jazyk $L_n$. Automat uvádzame pomocou diagramu.




Obr. 2.7: automat $A_n$ pre jazyk $\{b\}.\{w \in \{a,b\}^* | \#_a(w) = n\}$

Uvažujme množinu dvojíc slov $F_n = \{(\varepsilon, ba^n)\} \cup \{(ba^k, a^{n-k} | 0 \leq k \leq n)\}$. Množina $F_n$ je podla definície 1.4.2 rozšírenou mätúcou množinou pre jazyk $L_n$. Nakoľko $|F_n| = n + 2$, tak podľa Vety 1.4.2 $nsc(L_n) \geq n + 2$. Nakoľko $L(A_n) = L_n$ a $\#_S(A) = n + 2$, tak $nsc(L_n) = n+2$ a automat $A_n$ je minimálnym NKA pre jazyk $L_n$. Teraz zostrojíme netriviálny rozklad automatu $A_n$. Rozklad uvádzame pomocou diagramu.




Obr. 2.8: netriviálny rozklad automatu $A_n$ pre jazyk $\{b\}.\{w \in \{a,b\}^* | \#_a(w) = n\}$ na automaty $A_1^n$(hore) a $A_2^n$(dole)

$L(A_1^n) = \{w \in \{a,b\}^* | \#_a(w) = n\}$, $L(A_2^n) = \{b\}.\{a,b\}^*$. Teda $L(A_1^n) \cap L(A_2^n) = L(A_n)$. Nakoľko $\#_S(A_1^n) < \#_S(A_n)$ a $\#_S(A_2^n) < \#_S(A_n)$, tak automaty $A_1^n$ a $A_2^n$ tvoria netriviálny rozklad automatu $A_n$. $\square$

**Tvrdenie 2.1.4.** *Pre $n, m \geq 2, 0 \leq z_n < n, 0 \leq z_m < m$ definujeme $L[n,m,z_n,z_m] = \{w \in \{a,b\}^* | \#_a(w) \equiv z_n (mod\, n),\ \#_b(w) \equiv z_m (mod\, m)\}$. Platia nasledovné tvrdenia:*

*(a) Jazyk $L[n,m,z_n,z_m]$ je rozložiteľný.*

*(b) Jazyk $L[n,m,z_n,z_m] \cup \{\varepsilon\}$ je rozložiteľný.*

*Dôkaz.* Najprv dokážeme (a). Uvažujme $n, m \geq 2$. Najprv ukážeme, že $nsc(L[n,m,z_n,z_m]) = nm$. Definujme NKA $A[n,m,z_n,z_m] = (K_A, \{a,b\}, \delta_A, q[0,0], \{q[z_n,z_m]\})$, kde $K_A = \{q[i,j] \mid 0 \leq i < n,\ 0 \leq j < m\}$ a prechodová funkcia $\delta_A$ je pre $0 \leq i < n,\ 0 \leq j < m$ definovaná nasledovne: $\delta_A(q[i,j],a) = \{q[(i+1)\ mod\ n, j]\}$, $\delta_A(q[i,j],b) = \{q[i,(j+1)\ mod\ m]\}$. Dá sa lahko nahliadnuť, že $L(A[n,m,z_n,z_m]) = L[n,m,z_n,z_m]$. Teraz uvažujme množinu dvojíc slov $S = \{(a^l b^k, a^{z_n+n-l} b^{z_m+m-k}) \mid 0 \leq l < n,\ 0 \leq k < m\}$. Množina $S$ je podla definície 1.4.1 mätúcou množinou pre jazyk $L[n,m,z_n,z_m]$. Kedže $|S| =$

$nm$, tak podľa Vety 1.4.1 platí $nsc(L[n,m,z_n,z_m]) \geq nm$. Nakoľko $L(A[n,m,z_n,z_m]) = L[n,m,z_n,z_m]$ a $\#_S(A[n,m,z_n,z_m]) = nm$, tak $nsc(L[n,m,z_n,z_m]) = nm$ a automat $A[n,m,z_n,z_m]$ je minimálny NKA pre jazyk $L[n,m,z_n,z_m]$.

Teraz zostrojíme netriviálny rozklad automatu $A[n,m,z_n,z_m]$, čím skompletizujeme dôkaz. Uvažujme NKA definované nasledovne:

1. $A[n,z_n] = (K[n,z_n], \{a,b\}, \delta[n,z_n], q[0], \{q[z_n]\})$ kde $K[n,z_n] = \{q[i] | 0 \leq i < n\}$ a prechodová funkcia $\delta[n,z_n]$ je pre $0 \leq i < n$ definovaná nasledovne: $\delta[n,z_n](q[i],a) = \{q[(i+1)\ mod\ n]\}$, $\delta[n,z_n](q[i],b) = \{q[i]\}$.

2. $A[m,z_m] = (K[m,z_m], \{a,b\}, \delta[m,z_m], q[0], \{q[z_m]\})$ kde $K[m,z_m] = \{q[i] | 0 \leq i < m\}$ a prechodová funkcia $\delta[m,z_m]$ je pre $0 \leq i < m$ definovaná nasledovne: $\delta[m,z_m](q_i,b) = \{q[(i+1)\ mod\ m]\}$, $\delta[m,z_m](q_i,a) = \{q[i]\}$.

Ľahko vidno, že $L(A[n,z_n]) = \{w \in \{a,b\}^* | \#_a(w) \equiv z_n (mod\ n)\}$, $L(A[m,z_m]) = \{w \in \{a,b\}^* | \#_b(w) \equiv z_m (mod\ m)\}$, teda $L(A[n,z_n]) \cap L(A[m,z_m]) = L(A[n,m,z_n,z_m])$. Nakoľko $\#_S(A[n,z_n]) < \#_S(A[n,m,z_n,z_m])$ a $\#_S(A[m,z_m]) < \#_S(A[n,m,z_n,z_m])$, tak automaty $A[n,z_n]$ a $A[m,z_m]$ tvoria netriviálny rozklad automatu $A[n,m,z_n,z_m]$.

Dokážeme (b). Ak $z_n = z_m = 0$, tak nie je čo dokazovať, nakoľko potom $L[n,m,z_n,z_m] \cup \{\varepsilon\} = L[n,m,z_n,z_m]$. Uvažujme teda $z_n \neq 0$ alebo $z_m \neq 0$. Uvažujme automat $A[n,m,z_n,z_m]$ z dôkazu (a) a na jeho základe definujme NKA $A_\varepsilon[n,m,z_n,z_m] = (K_A \cup \{q_\varepsilon\}, \{a,b\}, \delta_\varepsilon, q_\varepsilon, \{q_\varepsilon, q[z_n,z_m]\})$ kde prechodová funkcia $\delta_\varepsilon$ je definovaná nasledovne: $\delta_\varepsilon(q_\varepsilon, \varepsilon) = \{q[0,0]\}$, $\forall q \in K_A, x \in \{a,b\} : \delta_\varepsilon(q,x) = \delta_A(q,x)$. Dá sa ľahko nahliadnuť, že $L(A_\varepsilon[n,m,z_n,z_m]) = L[n,m,z_n,z_m] \cup \{\varepsilon\}$. Uvažujme množinu dvojíc slov $M_\varepsilon = \{(a^{l_n}b^{l_m}, a^{n+z_n-l_n}b^{m+z_m-l_m})\} \cup \{(\varepsilon,\varepsilon)\}$. Množina $M_\varepsilon$ je podľa definície 1.4.2 rozšírenou mätúcou množinou pre jazyk $L[n,m,z_n,z_m] \cup \{\varepsilon\}$. Kedže $|M_\varepsilon| = nm+1$, tak podľa Vety 1.4.2 platí $nsc(L[n,m,z_n,z_m] \cup \{\varepsilon\}) \geq nm+1$. Nakoľko $L(A_\varepsilon[n,m,z_n,z_m]) = L[n,m,z_n,z_m] \cup \{\varepsilon\}$ a $\#_S(A_\varepsilon[n,m,z_n,z_m]) = nm+1$, tak $nsc(L[n,m,z_n,z_m] \cup \{\varepsilon\}) = nm+1$ a automat $A_\varepsilon[n,m,z_n,z_m]$ je minimálny NKA pre jazyk $L[n,m,z_n,z_m] \cup \{\varepsilon\}$.

Zostrojíme netriviálny rozklad automatu $A_\varepsilon[n,m,z_n,z_m]$ čím skompletizujeme dôkaz. Uvažujme NKA definované nasledovne:

1. Na základe automatu $A[n,z_n]$ z dôkazu (a) definujeme NKA $A_\varepsilon[n,z_n] = (K[n,z_n] \cup \{q_\varepsilon\}, \{a,b\}, \delta_\varepsilon[n,z_n], q_\varepsilon, \{q_\varepsilon, q[z_n]\})$ kde prechodová funkcia $\delta_\varepsilon[n,z_n]$ je definovaná nasledovne: $\delta_\varepsilon[n,z_n](q_\varepsilon,\varepsilon) = \{q[0]\}, \forall q \in K[n,z_n], x \in \{a,b\} : \delta_\varepsilon[n,z_n](q,x) = \delta[n,z_n](q,x)$.

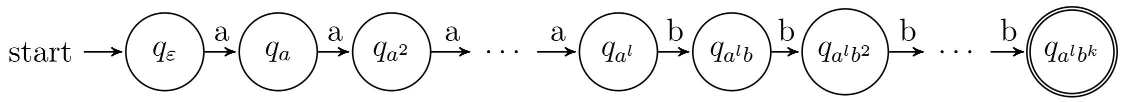
2. Na základe automatu $A[m,z_m]$ z dôkazu (a) definujeme NKA $A_\varepsilon[m,z_m] = (K[m,z_m] \cup \{q_\varepsilon\}, \{a,b\}, \delta_\varepsilon[m,z_m], q_\varepsilon, \{q_\varepsilon, q[z_m]\})$ kde prechodová funkcia $\delta_\varepsilon[m,z_m]$ je definovaná nasledovne: $\delta_\varepsilon[m,z_m](q_\varepsilon,\varepsilon) = \{q[0]\}, \forall q \in K[m,z_m], x \in \{a,b\} : \delta_\varepsilon[m,z_m](q,x) = \delta[m,z_m](q,x)$.

Ľahko vidno, že $L(A_\varepsilon[n, z_n]) = \{w \in \{a,b\}^* | \#_a(w) \equiv z_n(mod\ n)\} \cup \{\varepsilon\}$, $L(A_\varepsilon[m, z_m]) = \{w \in \{a,b\}^* | \#_b(w) \equiv z_m(mod\ m)\} \cup \{\varepsilon\}$, teda $L(A_\varepsilon[n, z_n]) \cap L(A_\varepsilon[m, z_m]) = L(A_\varepsilon[n, m, z_n, z_m])$. Nakoľko $\#_S(A_\varepsilon[n, z_n]) < \#_S(A_\varepsilon[n, m, z_n, z_m])$ a $\#_S(A_\varepsilon[m, z_m]) < \#_S(A_\varepsilon[n, m, z_n, z_m])$, tak automaty $A_\varepsilon[n, z_n]$ a $A_\varepsilon[m, z_m]$ tvoria netriviálny rozklad automatu $A_\varepsilon[n, m, z_n, z_m]$.

□

**Tvrdenie 2.1.5.** *Uvažujme jazyky $L_{l,k} = \{a^l b^k\}$. Ak $l, k \geq 1$, potom je jazyk $L_{l,k}$ rozložiteľný.*

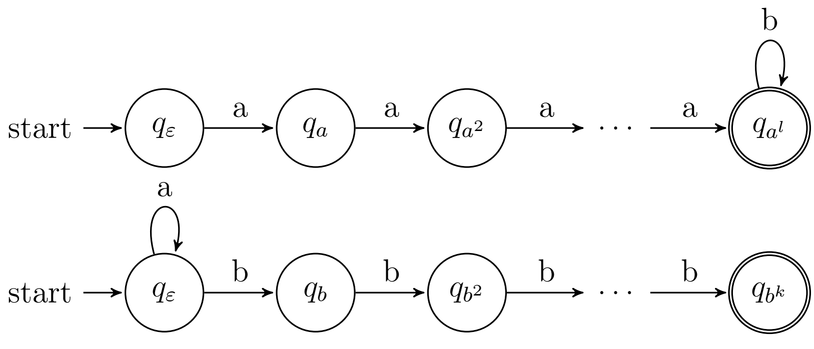
*Dôkaz.* Uvažujme $l, k \geq 1$. Ukážeme, že $nsc(L_{l,k}) = l + k + 1$. Najprv zostrojíme NKA $A_{l,k}$ pre jazyk $L_{l,k}$. Automat uvádzame pomocou diagramu.




Obr. 2.9: automat $A_{l,k}$ pre jazyk $\{a^l b^k\}$

Teraz uvažujme množinu dvojíc slov $F = \{(a^i, a^{l-i} b^k), (a^l b^j, b^{k-j}) | 0 \leq i \leq l,\ 1 \leq j \leq k\}$. Množina $F$ je podľa definície 1.4.1 mätúcou množinou pre jazyk $L_{l,k}$. Keďže $|F| = l + k + 1$, tak podľa Vety 1.4.1 platí $nsc(L_{l,k}) \geq l + k + 1$. Nakoľko $L(A_{l,k}) = L_{l,k}$ a $\#_S(A_{l,k}) = l + k + 1$, tak $nsc(L_{l,k}) = l + k + 1$ a automat $A_{l,k}$ je minimálnym NKA pre jazyk $L_{l,k}$.

Teraz zostrojíme netriviálny rozklad automatu $A_{l,k}$. Hľadané automaty $A_l$ a $A_k$ uvádzame pomocou diagramov.




Obr. 2.10: rozklad automat $A_{l,k}$ na automaty $A_l$(hore) a $A_k$(dole)

Ľahko vidno, že $L(A_l) = \{a^l b^i | i \in \mathbb{N}\}$ a $L(A_k) = \{a^i b^k | i \in \mathbb{N}\}$. Teda $L(A_l) \cap L(A_k) = L(A_{l,k})$. Navyše $\#_S(A_l) < \#_S(A_{l,k})$ a $\#_S(A_k) < \#_S(A_{l,k})$, teda automaty $A_l$ a $A_k$ tvoria netriviálny rozklad automatu $A_{l,k}$.

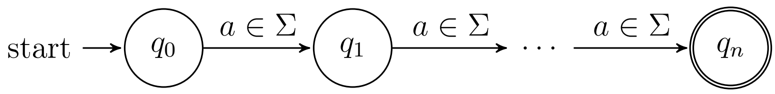
□

**Dôsledok 2.1.1.** *Existuje konečný jazyk, ktorý je rozložiteľný.*

## 2.2 Nerozložiteľné jazyky

**Tvrdenie 2.2.1.** *Pre ľubovolnú abecedu $\Sigma$ a každé $n \in \mathbb{N}$ je jazyk $\Sigma^n$ nerozložiteľný.*

*Dôkaz.* Uvažujeme $n \in \mathbb{N}$. Najprv ukážeme, že $nsc(\Sigma^n) = n + 1$. Najprv zostrojme NKA $A_{\Sigma^n}$ taký, že $L(A_{\Sigma^n}) = \Sigma^n$. Automat uvádzame pomocou diagramu.




Obr. 2.11: automat $A_{\Sigma^n}$

Vezmime ľubovolné $a \in \Sigma$ a uvažujme množinu $F = \{(a^i, a^{n-i}) \mid 0 \leq i \leq n\}$. Množina $F$ je podľa definície 1.4.1 mätúcou množinou pre jazyk $\Sigma^n$, teda podľa Vety 1.4.1 platí $nsc(\Sigma^n) \geq n + 1$. Nakoľko sme zostrojili NKA akceptujúci jazyk $\Sigma^n$, ktorý má práve $n + 1$ stavov, tak $nsc(\Sigma^n) = n + 1$ a NKA $A_{\Sigma^n}$ je minimálnym automatom pre jazyk $\Sigma^n$.

Pre $n = 0$ a $n = 1$ vyplýva platnosť tvrdenia z Vety 4.1.1. Pre $n \geq 2$ postupujme sporom. Nech je jazyk $\Sigma^n$ rozložiteľný, teda existuje netriviálny rozklad automatu $A_{\Sigma^n}$. To znamená, že existujú NKA $A_1^{\Sigma^n}$ a $A_2^{\Sigma^n}$ také, že $L(A_1^{\Sigma^n}) \cap L(A_2^{\Sigma^n}) = \Sigma^n$ a $\#_S(A_1^{\Sigma^n}) < n+1$, $\#_S(A_2^{\Sigma^n}) < n+1$. Navyše vďaka Leme 1.3.1 môžeme predpokladať, že automaty $A_1^{\Sigma^n}$ a $A_2^{\Sigma^n}$ neobsahujú prechody na $\varepsilon$.

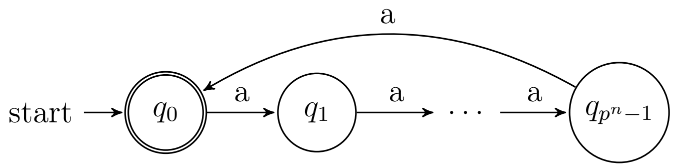
Vezmime ľubovoľné $a \in \Sigma$ a uvažujme výpočet automatu $A_1^{\Sigma^n}$ na slove $a^n$. Podľa predchádzajúceho automat $A_1^{\Sigma^n}$ slovo $a^n$ akceptuje. Výpočet vyzerá nasledovne: $(p_0, a^n) \vdash (p_1, a^{n-1}) \vdash \ldots \vdash (p_{n-1}, a) \vdash (p_n, \varepsilon)$ kde $p_0 = q_0$, $A_1^{\Sigma^n}$, $p_n \in F_{A_1^{\Sigma^n}}$ a pre $1 \leq i < n$ $p_i \in K_{A_1^{\Sigma^n}}$. Nakoľko $\#_S(A_1^{\Sigma^n}) < n + 1$, tak $\exists i, j \in \mathbb{N} : 0 \leq i \leq, ni \neq j, p_i = p_j$ (vo výpočte sa nejaký stav zopakuje). Z toho vyplýva, že v akceptovanom slove môžem nejakú jeho časť pumpovať, t.j. $\exists r_1 \in \mathbb{N}, 1 \leq r_1 \leq n\ \forall k \in \mathbb{N} : a^{n+kr_1} \in L(A_1^{\Sigma^n})$.

Analogicky, uvažujúc výpočet automatu $A_2^{\Sigma^n}$ na slove $a^n$, platí $\exists r_2 \in \mathbb{N}, 1 \leq r_2 \leq n\ \forall k \in \mathbb{N} : a^{n+kr_2} \in L(A_2^{\Sigma^n})$.

Teraz uvažujme slovo $a^{n+r_1r_2}$. Podľa predchádzajúceho platí $a^{n+r_1r_2} \in L(A_1^{\Sigma^n}) \cap L(A_2^{\Sigma^n})$. Avšak $a^{n+r_1r_2} \notin \Sigma^n$ čo je v spore s tým, že automaty $A_1^{\Sigma^n}$ a $A_2^{\Sigma^n}$ tvoria netriviálny rozklad automatu $A_{\Sigma^n}$. $\square$

**Veta 2.2.1.** *Pre $n \geq 1$ a $p$ je prvočíslo definujeme $L_{p^n} = \{a^{kp^n} | k \in \mathbb{N}\}$. Jazyk $L_{p^n}$ nerozložiteľný.*

*Dôkaz.* Najprv ukážeme, že $nsc(L_{p^n}) = p^n$. Zostrojme NKA $A_{p^n}$ taký, že $L(A_{p^n}) = L_{p^n}$. Automat uvádzame pomocou diagramu.




Obr. 2.12: automat $A_{p^n}$

Uvažujme množinu dvojíc slov $F = \{(a^l, a^{p^n-l}) \mid 0 \leq l \leq p^n - 1\}$. Množina $F$ je podľa definície 1.4.1 mätúcou množinou pre jazyk $L_{p^n}$. Nakoľko $|F| = p^n$, tak podľa Vety 1.4.1 platí $nsc(L_{p^n}) \geq p^n$. Keďže sa nám podarilo zostrojiť automat akceptujúci $L_{p^n}$, ktorý má práve $p^n$ stavov, tak platí $nsc(L_{p^n}) = p^n$. Intuitívne je jasné, že automat „počíta zvyšok po delení $p^n$“.

Ďalej postupujme sporom. Uvažujme, že jazyk $L_{p^n}$ je rozložiteľný, teda že existuje netriviálny rozklad automatu $A_{p^n}$. To znamená, že existujú NKA $A_1^{p^n}, A_2^{p^n}$, také, že platí $\#_S(A_1^{p^n}) < p^n$, $\#_S(A_2^{p^n}) < p^n$, $L(A_1^{p^n}) \cap L(A_2^{p^n}) = L_{p^n}$. Navyše podľa Lemy 1.3.1 môžeme predpokladať, že automaty $A_1^{p^n}$ a $A_2^{p^n}$ neobsahujú prechody na $\varepsilon$.

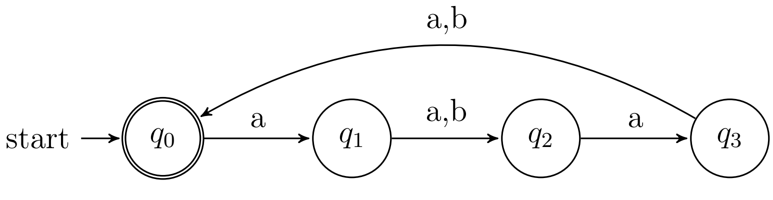
Z predchádzajúceho vyplýva, že $a^{p^n} \in L(A_1^{p^n}), a^{p^n} \in L(A_2^{p^n})$. Teraz sa pozrime na výpočet automatu $A_1^{p^n}$ na slove $a^{p^n}$. Nech tento výpočet vyzerá nasledovne $(q_0, a^{p^n}) \vdash (q_1, a^{p^n-1}) \vdash \cdots \vdash (q_{p^n-1}, a) \vdash (q_{p^n}, \varepsilon)$, kde $q_0$ je počiatočný stav automatu $A_1^{p^n}$, $q_{p^n}$ je nejaký akceptačný stav automatu $A_1^{p^n}$ a pre $1 \leq i < p^n$ $q_i \in K_{A_1^{p^n}}$. Nakoľko $\#_S(A_1^{p^n}) < p^n$, tak nutne $\exists i, j \in \mathbb{N}, 0 \leq i, j < p^n, i \neq j : q_i = q_j$ (počas výpočtu sa v časti „od začiatku po predposledný stav“ nejaký stav zopakuje). Z toho vyplýva, že v akceptovanom slove môžem pumpovať časť, ktorá je kratšia ako $p^n$, t.j. $\exists r_1 \in \mathbb{N}, 1 \leq r_1 < p^n\ \forall k \in \mathbb{N} : a^{p^n+kr_1} \in L(A_1^{p^n})$.

Analogicky, uvažujúc výpočet automatu $A_2^{p^n}$ na slove $a^{p^n}$, platí $\exists r_2 \in \mathbb{N}, 1 \leq r_2 < p^n\ \forall k \in \mathbb{N} : a^{p^n+kr_2} \in L(A_2^{p^n})$.

Čísla $r_1$ a $r_2$ zapíšme nasledovne. $r_1 = p^{l_1}s_1, 0 \leq l_1 < n, p \nmid s_1$. $r_2 = p^{l_2}s_2, 0 \leq l_2 < n, p \nmid s_2$. Z uvedeného v predošlom vyplýva, že $a^{p^n+p^{max(l_1,l_2)}s_1s_2} \in L(A_1^{p^n}) \cap L(A_2^{p^n})$. Nakoľko však $p^n \nmid p^{max(l_1,l_2)}s_1s_2$, tak $a^{p^n+p^{max(l_1,l_2)}s_1s_2} \notin L_{p^n}$, čo je však v spore s predpokladom, že automaty $A_1^{p^n}$ a $A_2^{p^n}$ tvoria netriviálny rozklad automatu $A_{p^n}$. □

**Tvrdenie 2.2.2.** *Jazyk* $L = (\{a\}\{a,b\}\{a\}\{a,b\})^*$ *je nerozložiteľný.*

*Dôkaz.* Najprv ukážeme, že $nsc(L) = 4$. Zostrojme NKA $A_L$ taký, že $L(A_L) = L$. Automat uvádzame pomocou diagramu.




Obr. 2.13: automat $A_L$ pre jazyk $L = (\{a\}\{a,b\}\{a\}\{a,b\})^*$

Uvažujme množinu $F = \{(\varepsilon, aaaa), (a, aaa), (aa, aa), (aaa, a)\}$. Množina $F$ je podľa definície 1.4.1 mätúcou množinou pre jazyk $L$, teda podľa Vety 1.4.1 platí $nsc(L) \geq 4$. Nakoľko sme zostrojili NKA akceptujúci jazyk $L$, ktorý má práve 4 stavy, tak $nsc(L) = 4$ a NKA $A_L$ je minimálnym automatom pre jazyk $L$.

Nech je jazyk $L$ rozložiteľný, teda existuje netriviálny rozklad automatu $A_L$. To znamená, že existujú NKA $A_1^L$ a $A_2^L$ také, že $L(A_1^L) \cap L(A_2^L) = L$ a $\#_S(A_1^L) < 4$, $\#_S(A_2^L) < 4$. Navyše vďaka Leme 1.3.1 môžeme predpokladať, že automaty $A_1^L$ a $A_2^L$ neobsahujú prechody na $\varepsilon$.

Uvažujme výpočet automatu $A_1^L$ na slove $aaaa$. Podľa predcházajúceho automat $A_1^L$ slovo $aaaa$ akceptuje. Výpočet vyzerá nasledovne: $(p_0, aaaa) \vdash (p_1, aaa) \vdash (p_2, aa) \vdash (p_3, a) \vdash (p_4, \varepsilon)$ kde $p_0$ je počiatočný stav $A_1^L$, $p_4 \in F_{A_1^L}$ a pre $1 \leq i < 4$ $p_i \in K_{A_1^L}$. Nakoľko $\#_S(A_1^L) < 4$, tak $\exists i, j \in \{0, 1, 2, 3\}, i \neq j, p_i = p_j$ (vo výpočte sa nejaký stav zopakuje ešte pred tým ako bude slovo akceptované) . Z toho vyplýva, že v akceptovanom slove môžem nejakú jeho časť pumpovať, t.j. $\exists r_1 \in \{1, 2, 3\}\ \forall k \in \mathbb{N} : a^{4+kr_1} \in L(A_1^L)$.

Analogicky, uvažujúc výpočet automatu $A_2^L$ na slove $aaaa$, platí $\exists r_2 \in \{1, 2, 3\}\ \forall k \in \mathbb{N} : a^{n+kr_2} \in L(A_2^L)$.

Môžu nastať nasledovné prípady:

1. $r_1 = 1, r_2 = 3$ respektíve $r_1 = 3, r_2 = 1$. V tom prípade podľa predchádzajúceho platí $a^{4+3} \in L(A_1^L) \cap L(A_2^L)$. Avšak $a^{4+3} \notin L$ čo je v spore s tým, že automaty $A_1^L$ a $A_2^L$ tvoria netriviálny rozklad automatu $A_L$.

2. $r_1 = 2, r_2 = 2$. V tom prípade podľa predchádzajúceho platí $a^{4+2} \in L(A_1^L) \cap L(A_2^L)$. Avšak $a^{4+2} \notin L$ čo je v spore s tým, že automaty $A_1^L$ a $A_2^L$ tvoria netriviálny rozklad automatu $A_L$.

Nakoľko iné prípady nastať nemôžu, našli sme hľadaný spor čo kompletizuje dôkaz. $\square$

# Kapitola 3

# Porovnanie deterministickej a nedeterministickej rozložiteľnosti regulárnych jazykov

Zaujímavou otázkou je, či existuje regulárny jazyk taký, že je deterministicky nerozložiteľný a súčasne nedeterminiticky rozložiteľný respektíve deterministicky rozložiteľný a súčasne nedeterminiticky nerozložiteľný.

## 3.1 Definícia deterministického konečného automatu

Pred tým, ako uvedieme dosiahnuté výsledky zavedieme definíciu deterministického konečného automatu, ktorú budeme používať, nakoľko existuje viacero prístupov k definovaniu deterministických konečných automatov.

**Definícia 3.1.1.** ***Deterministický konečný automat*** *je pätica* $(K, \Sigma, \delta, q_0, F)$*, kde:*

1. *$K$ je konečná množina stavov*
2. *$\Sigma$ je konečná vstupná abeceda*
3. *$q_0 \in K$ je počiatočný stav*
4. *$F \subseteq K$ je množina akceptačných stavov*
5. *$\delta : K \times \Sigma \to K$ je prechodová funkcia*

**Poznámka 3.1.1.** *Deterministický konečný automat sa skrátene označuje DKA.*

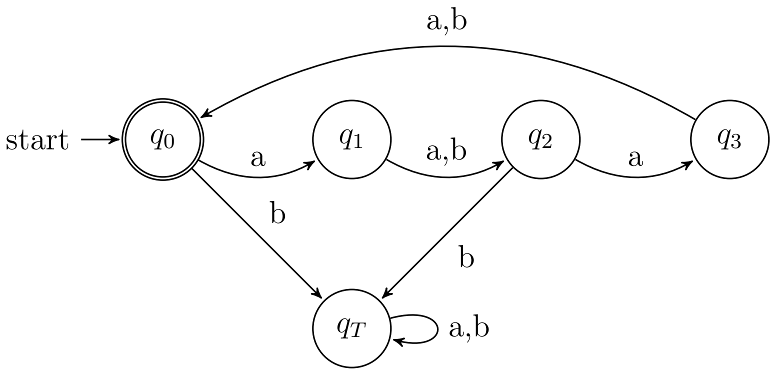
Poznajúc ako v našom texte definujeme deterministický konečný automat je pre čitateľa so základnými znalosťami v oblasti jasné, ako by boli definované ostatné potrebné pojmy (konfigurácia, krok výpočtu, akceptovaný jazyk, rozlozložiteľnosť DKA, deterministická rozložiteľnosť regulárneho jazyka), preto ich definície neuvádzame.

## 3.2 Rozdielové jazyky

Uvádzame rozdielové jazyky, ktoré ukazujú, že pojem rozložiteľnosti regulárneho jazyka je rôzny ak uvažujeme deterministické resp. nedeterministické automaty. Intuícia našepkáva, že ak to pôjde, tak by to mohlo ísť skôr tak, že nájdeme jazyk ktorý je deterministicky nerozložiteľný a zároveň nedeterministicky rozložiteľný (očakávali sme, že v rozklade ušetrí nedeterminizmus stavy). Avšak podarilo sa nám nájsť rozdielové jazyky, pri ktorých to je opačne. Prvým našim výsledkom v tomto smere je nasledujúce tvrdenie.

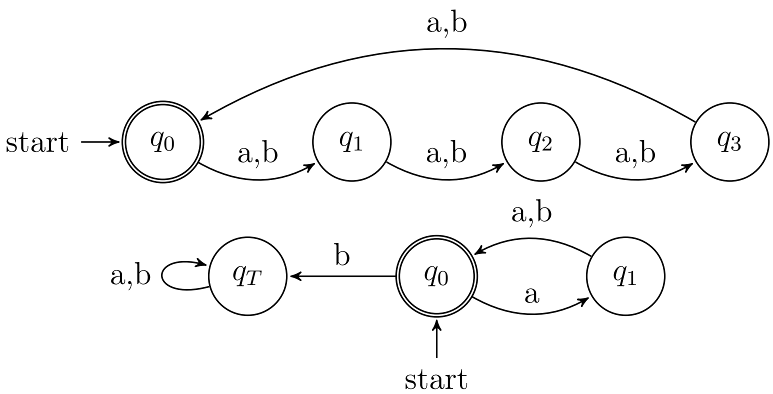
**Veta 3.2.1.** *Existuje nedeterministicky nerozložiteľný deterministicky rozložiteľný regulárny jazyk.*

*Dôkaz.* Hľadaným jazykom je jazyk $L = (\{a\}\{a,b\}\{a\}\{a,b\})^*$. Ukážeme, že jazyk $L$ je deterministicky rozložiteľný. Najprv zostrojíme minimálny DKA $A_L$ akceptujúci $L$. Automat uvádzame pomocou diagramu.




Obr. 3.1: DKA $A_L$ pre jazyk $L = (\{a\}\{a,b\}\{a\}\{a,b\})^*$

Ľahko vidno, že $A_L$ akceptuje práve $L$. Minimalita $A_L$ sa dá dokázať pomocou Myhill-Nerodeovej vety. Zostrojíme netriviálny rozklad automatu $A_L$. Hľadané DKA $A_1^L$ a $A_2^L$ uvádzame pomocou ich diagramov.




Obr. 3.2: rozklad automatu $A_L$ na automaty $A_1^L$(hore) a $A_2^L$(dole)

Možno nahliadnuť, že jeden z automatov v rozklade počíta zvyšok po delení 4 a druhý kontroluje, či symboly na nepárnych pozíciách v slove sú $a$. Teda vidno, že $L(A_1^L) = \{w \in \{a,b\}^* \mid |w| \equiv 0 \ (mod\ 4)\}$ a $L(A_2^L) = (\{a\}\{a,b\})^*$. Teda $L(A_1^L) \cap L(A_2^L) = L$. Navyše $\#_S(A_1^L) < \#_S(A_L)$ a $\#_S(A_2^L) < \#_S(A_L)$, teda automaty $A_1^L$ a $A_2^L$ tvoria netriviálny rozklad automatu $A_L$. Z predchádzajúceho vyplýva, že jazyk $L = (\{a\}\{a,b\}\{a\}\{a,b\})^*$ je deterministicky rozložiteľný. Avšak tento jazyk je podľa Tvrdenia 2.2.2 nedeterministicky nerozložiteľný. □

Uvedená Veta síce ukazuje rozdiel medzi deterministickou a nedeterministickou rozložiteľnosťou, avšak jej dôkaz veľmi závisí od faktu, že v definícii DKA požadujeme úplnú prechodovú funkciu, vďaka čomu DKA použitý v dôkaze musí mať odpadový stav. Bez tohto odpadového stavu by náš dôkaz neprešiel. Nasledujúca Veta ukazuje, že existujú prípady, kde rozdiel medzi deterministickou a nedeterministickou rozložiteľnosťou nie je spôsobený iba nutnosťou úplnej prechodovej funkcie DKA.

**Veta 3.2.2.** *Existuje postupnosť jazykov $(L_i)_{i=2}^{\infty}$, taká, že platí:*

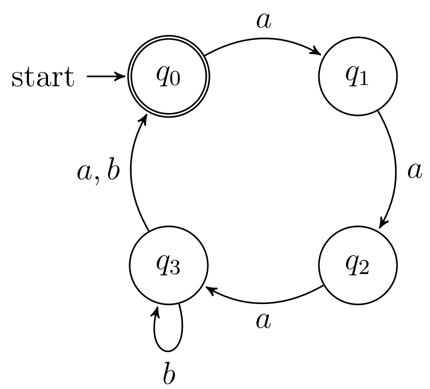
*(a) Jazyk $L_i$ je nedeterministicky nerozložiteľný a súčasne deterministicky rozložiteľný pre ľubovolné $i \in \mathbb{N}, i \geq 2$.*

*(b) Nech pre ľubovolné $i \in \mathbb{N}, i \geq 3$ je $A_i$ minimálny DKA akceptujúci $L_i$. Potom existuje taký rozklad $A_i$ na $A_1^i$ a $A_2^i$, že platí $\#_S(A_1^i) = \#_S(A_2^i) = \frac{\#_S(A_i)+3}{2}$.*

*Dôkaz.* Definujme postupnosť jazykov $(L_i')_{i=2}^{\infty}$ nasledovne: $L_i' = (\{a^{i-1}\}\{b\}^*\{a,b\})^*$ pre ľubovolné $i \geq 2$. Hľadanú postupnosť $(L_i)_{i=2}^{\infty}$ dostaneme z $(L_i')_{i=2}^{\infty}$ vybratím niektorých (nekonečne veľa) jej členov.

Najskôr ukážeme, že pre spočítateľne veľa $i \geq 2$ je $L_i'$ nedeterministicky nerozložiteľný. V nasledujúcom uvažujme teda $i \geq 2$ také, že $i$ je mocninou prvočísla. Zostrojíme NKA $A_i^N$ taký, že $L(A_i^N) = L_i'$. $A_i^N = (K_i^N, \{a,b\}, \delta_i^N, q_0, \{q_0\})$. Kde $K_i^N = \{q_j \mid 0 \leq j < i\}$ a prechodová funkcia $\delta_i^N$ je definovaná nasledovne - $\forall 0 \leq j \leq i-2 : \delta_i^N(q_j, a) = \{q_{j+1}\}$, $\delta_i^N(q_{i-1}, a) = \{q_0\}$, $\delta_i^N(q_{i-1}, b) = \{q_0, q_{i-1}\}$

Pre ilustráciu a lepšiu čitateľnosť uvádzame automat $A_4^N$ aj pomocou diagramu.




Obr. 3.3: automat $A_4^N$

Dá sa nahliadnuť, že platí $L(A_i^N) = L'_i$. Navyše, je dobré uvedomiť si, že $\{a^{ki} \mid i \in \mathbb{N}\} \subset L'_i$. Uvažujme množinu dvojíc slov $M_i = \{(a^j, a^{i-j}) \mid 0 \leq j < i\}$. Podľa definície 1.4.1 je množina $M_i$ rozšírenou mätúcou množinou pre jazyk $L'_i$. $|M_i| = i$, teda podľa Vety 1.4.1 $nsc(L'_i) \geq i$. Kedže $L(A_i^N) = L'_i$ a $\#_S(A_i^N) = i$, tak $nsc(L'_i) = i$ a automat $A_i^N$ je minimálny NKA pre jazyk $L'_i$.

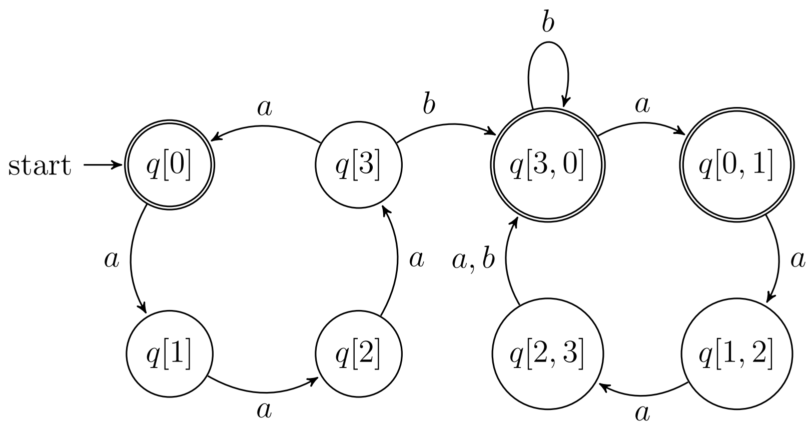
Ďalej postupujme sporom. Uvažujme, že jazyk $L_i$ je rozložiteľný, teda že existuje netriviálny rozklad automatu $A_i^N$. To znamená, že existujú NKA $A_1^{N,i}, A_2^{N,i}$, také, že platí $\#_S(A_1^{N,i}) < i$, $\#_S(A_2^{N,i}) < i$, $L(A_1^{N,i}) \cap L(A_2^{N,i}) = L'_i$. Navyše podľa Lemy 1.3.1 môžeme predpokladať, že automaty $A_1^{N,i}$ a $A_2^{N,i}$ neobsahujú prechody na $\varepsilon$.

Nakoľko $i$ je mocninou prvočísla, tak existuje nejaké prvočíslo $p$ a nejaké $n$také, že $i = p^n$. Z predchádzajúceho vyplýva, že $a^{p^n} \in L(A_1^{N,i}), a^{p^n} \in L(A_2^{N,i})$. Teraz sa pozrime na výpočet automatu $A_1^{N,i}$ na slove $a^{p^n}$. Nech tento výpočet vyzerá nasledovne $(q_0, a^{p^n}) \vdash (q_1, a^{p^n-1}) \vdash \cdots \vdash (q_{p^n-1}, a) \vdash (q_{p^n}, \varepsilon)$, kde $q_0$ je počiatočný stav automatu $A_1^{N,i}$, $q_{p^n}$ je nejaký akceptačný stav automatu $A_1^{N,i}$ a pre $1 \leq j < p^n$ $q_j \in K_{A_1^{N,i}}$. Nakoľko $\#_S(A_1^{N,i}) < p^n (= i)$, tak nutne $\exists j, k \in \mathbb{N}, 0 \leq j, k < p^n, j \neq k : q_j = q_k$ (počas výpočtu sa v časti „od začiatku po predposledný stav“ nejaký stav zopakuje). Z toho vyplýva, že v akceptovanom slove môžem pumpovať časť, ktorá je kratšia ako $p^n$, t.j. $\exists r_1 \in \mathbb{N}, 1 \leq r_1 < p^n \; \forall k \in \mathbb{N} : a^{p^n+kr_1} \in L(A_1^{p^n})$.

Analogicky, uvažujúc výpočet automatu $A_2^{N,i}$ na slove $a^{p^n}$, platí $\exists r_2 \in \mathbb{N}, 1 \leq r_2 < p^n \; \forall k \in \mathbb{N} : a^{p^n+kr_2} \in L(A_2^{N,i})$.

Čísla $r_1$ a $r_2$ zapíšme nasledovne. $r_1 = p^{l_1}s_1,\ 0 \leq l_1 < n,\ p \nmid s_1$. $r_2 = p^{l_2}s_2,\ 0 \leq l_2 < n,\ p \nmid s_2$. Z uvedeného v predošlom vyplýva, že $a^{p^n+p^{max(l_1,l_2)}s_1s_2} \in L(A_1^{N,i}) \cap L(A_2^{N,i})$. Nakoľko však $p^n \nmid p^{max(l_1,l_2)}s_1s_2$, tak $a^{p^n+p^{max(l_1,l_2)}s_1s_2} \notin L_i$, čo je však v spore s predpokladom, že automaty $A_1^{N,i}$ a $A_2^{N,i}$ tvoria netriviálny rozklad automatu $A_i^N$. Teda jazyk $L'_i$ je nedeterministicky nerozložiteľný.

Ukážeme, že $L'_i$ je deterministicky rozložiteľný pre ľubovoľné $i \geq 2$. DKA $A_i^D$ taký, že $L(A_i^D) = L'_i$, zostrojíme štandardnou podmnožinovou konštrukciou z NKA $A_i^D$. Takto dostaneme $A_i^D = (\{q[j], q[j,(j+1)\ mod\ i] \mid 0 \leq j < i\} \cup \{q_T\}, \{a, b\}, \delta_i^D, q[0], \{q[0], q[i-1, 0], q[0, 1]\})$, kde prechodová funkcia $\delta_i^D$ je definovaná nasledovne: $\delta_i^D(q[i-1], b) = q[i-1, 0]$, $\delta_i^D(q[i-1, 0], b) = q[i-1, 0]$, $\delta_i^D(q[i-2, i-1], b) = q[i-1, 0]$, $\forall j \in \mathbb{N}, 0 \leq j < i : \delta_i^D(q[j], a) = q[(j+1)\ mod\ i], \delta_i^D(q[j,(j+1)\ mod\ i], a) = q[(j+1)\ mod\ i, (((j+1)\ mod\ i) + 1)\ mod\ i]$. Naša definícia DKA požaduje úplnosť prechodovej funkcie, teda zatiaľ nedefinované prechody dodefinujeme tak, že idú automaticky do odpadového stavu $q_T$. Minimalita $A_i^D$ sa dá dokázať pomocou Myhill-Nerodeovej vety. Pre ilustráciu a lepšiu čitateľnosť uvádzame automat $A_4^D$ aj pomocou diagramu. V diagrame pre prehľadnosť neuvádzame odpadový stav $q_T$.

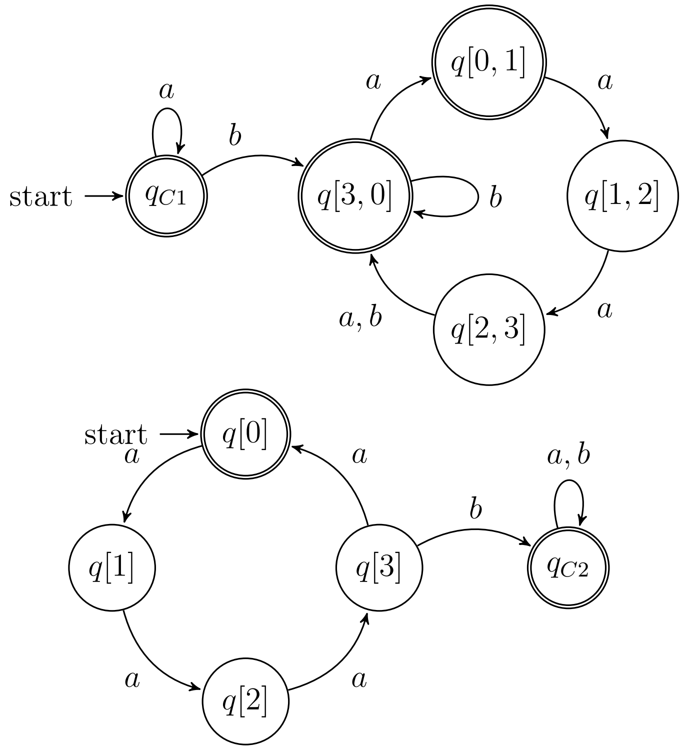



Obr. 3.4: automat $A_4^D$

Zostrojíme netriviálny rozklad DKA $A_i^D$. Myšlienkou rozkladu je, neformálne povedané, že v jednotlivých automatoch rozkladu budeme mať namiesto oboch úplných cyklov, ktoré sú v $A_i^D$ jeden cyklus „zlúčený“ do jedného stavu a druhý cyklus úplný. Keď budeme uvažovať slová, ktoré budú akceptované oboma automatmi, tak vždy jeden automat správne zráta daný cyklus, čo nám bude stačiť. Uvedomme si ešte, že rozklad nám bude správne fungovať aj vďaka faktu, že dané dva cykly sú oddelené práve jedným prechodom na $b$ z prvého cyklu do druhého, pričom v prvom cykle prechody na $b$ nepoužívame. Formálne definujeme automaty $A_1^{D,i}, A_2^{D,i}$. Pre ilustráciu a lepšiu čitateľnosť uvádzame rozklad automatu $A_4^D$ na automaty $A_1^{D,4}, A_2^{D,4}$ aj pomocou diagramu Obr. 3.2. V diagrame pre prehľadnosť neuvádzame odpadový stav $q_T$.

1. $A_1^{D,i} = (\{q_{C1}, q_T\} \cup K_1^{D,i}, \{a, b\}, \delta_1^{D,i}, q_{C1}, F_1^{D,i})$, kde $K_1^{D,i} = \{q[j, (j+1)\ mod\ i] \mid 0 \leq j < i\}$, $F_1^{D,i} = \{q_{C1}, q[i-1, 0], q[0, 1]\}$ a prechodová funkcia $\delta_1^{D,i}$ je definovaná nasledovne: $\delta_1^{D,i}(q_{C1}, a) = q_{C1}$, $\delta_1^{D,i}(q_{C1}, b) = q[i-1, 0]$, $\delta_i^D(q[i-1, 0], b) = q[i-1, 0]$, $\delta_i^D(q[i-2, i-1], b) = q[i-1, 0]$, $\forall q[j, k] \in K_1^{D,i} : \delta_i^D(q[j, k], a) = \{q[(j+1)\ mod\ i, (k+1)\ mod\ i]\}$. Prechody, ktoré sme zatiaľ nedefinovali, dodefinujeme tak, že automaticky vedú do odpadového stavu $q_T$

2. $A_2^{D,i} = (\{q_{C2}, q_T\} \cup \{q[j] \mid 0 \leq j < i\}, \{a, b\}, \delta_2^{D,i}, q[0], F_2^{D,i})$, kde $F_2^{D,i} = \{q[0], q_{C2}\}$ a prechodová funkcia $\delta_2^{D,i}$ je definovaná nasledovne: $\delta_2^{D,i}(q[i-1], b) = q_{C2}$, $\delta_1^{D,i}(q_{C2}, a) = q_{C2}$, $\delta_1^{D,i}(q_{C2}, b) = q_{C2}, \forall j \in \mathbb{N}, 0 \leq j < i : \delta_i^D(q[j], a) = q[(j+1)\ mod\ i]$. Prechody, ktoré sme zatiaľ nedefinovali, dodefinujeme tak, že automaticky vedú do odpadového stavu $q_T$




Obr. 3.5: rozklad automatu $A_4^D$ na automaty $A_1^{D,4}$(hore) a $A_2^{D,4}$(dole)

Ukážeme, že $L(A_i^D) = L(A_1^{D,i}) \cap L(A_2^{D,i})$.
$\subseteq$: Ľahko vidno z konštrukcie automatov $A_1^{D,i}$ a $A_2^{D,i}$.
$\supseteq$: Uvažujme $w \in L(A_1^{D,i}) \cap L(A_2^{D,i})$. Teda existujú stavy $q_{F1} \in F_1^{D,i}, q_{F2} \in F_2^{D,i}$ také, že $(q_{C1}, w) \vdash^*_{A_1^{D,i}} (q_{F1}, \varepsilon), (q[0], w) \vdash^*_{A_2^{D,i}} (q_{F2}, \varepsilon)$. Rozoberieme postupne, aký môže byť stav $q_{F1}$.

1. $q_{F1} = q_{C1}$. Z konštrukcie $A_1^{D,i}$ plynie, že výpočet $(q_{C1}, w) \vdash^*_{A_1^{D,i}} (q_{F1}, \varepsilon)$ prechádza iba cez stav $q_{C1}$. Teda existuje nejaké $n \in \mathbb{N}$ také, že $w = a^n$. Z konštrukcie $A_2^{D,i}$ teda $q_{F2} = q[0]$ a výpočet $(q[0], w) \vdash^*_{A_2^{D,i}} (q_{F2}, \varepsilon)$ je zároveň akceptačným výpočtom automatu $A_i^D$ na slove $w$. Neformálne, automat $A_i^D$ používa iba prvý svoj cyklus, ktorý je ale celý obsiahnutý aj v $A_2^{D,i}$.

2. $q_{F1} \in \{q[i-1, 0], q[0, 1]\}$. Potom z konštrukcie $A_1^{D,i}$ vyplýva, že existujú nejaké $n \in \mathbb{N}, u \in \{a, b\}^*$ také, že $w = a^n bu$. Výpočet automatu $A_1^{D,i}$ na slove $w$ teda vyzerá nasledovne: $(q_{C1}, a^n bu) \vdash^*_{A_1^{D,i}} (q_{C1}, bu) \vdash_{A_1^{D,i}} (q[i-1, 0], u) \vdash^*_{A_1^{D,i}} (q_{F1,\varepsilon})$. Nakoľko slovo $w$ obsahuje symbol $b$, tak z konštrukcie $A_2^{D,i}$ vyplýva $q_{F2} = q_{C2}$ a výpočet automatu $A_1^{D,i}$ na slove $w$ vyzerá nasledovne: $(q[0], a^n bu) \vdash^*_{A_2^{D,i}} (q[i-1], bu) \vdash_{A_2^{D,i}} (q_{C2}, u) \vdash^*_{A_2^{D,i}} (q_{C2}, \varepsilon)$. Z konštrukcie $A_1^{D,i}$ plynie, že výpočet $(q[i-1, 0], u) \vdash^*_{A_1^{D,i}} (q_{F1}, \varepsilon)$ v automate $A_1^{D,i}$ je zároveň výpočtom $(q[i-1, 0], u) \vdash^*_{A_i^D} (q_{F1}, \varepsilon)$ v automate $A_i^D$. Z konštrukcie $A_2^{D,i}$ plynie, že výpočet $(q[0], a^n bu) \vdash^*_{A_2^{D,i}} (q[i-1], bu)$ v automate $A_2^{D,i}$ je zároveň výpočtom $(q[0], a^n bu) \vdash^*_{A_i^D} (q[i-1], bu)$ v automate

$A_i^D$. Navyše platí $\delta_i^D(q[i-1], b) = q[i-1, 0]$. Z toho vyplýva $(q[0], a^n bu) \vdash^*_{A_i^D} (q[i-1], bu) \vdash_{A_i^D} (q[i-1,0], u) \vdash^*_{A_i^D} (q_{F1,\varepsilon})$. Nakoľko $q_{F1} \in F_i^D$, tak tento výpočet je akceptačným výpočtom automatu $A_i^D$ na slove $w$. Neformálne povedané, oba z automatov v rozklade zrátajú jeden cyklus z pôvodného automatu a v tom druhom len stoja. V prieniku dostaneme teda zrátané oba cykly.

Nakoľko iné možnosti neexistujú, tak platí $L(A_i^D) = L(A_1^{D,i}) \cap L(A_2^{D,i})$. Zjavne $\#_S(A_1^{D,i}) < \#_S(A_i^D), \#_S(A_2^{D,i}) < \#_S(A_i^D)$ a teda automaty $A_1^{D,i}$ a $A_2^{D,i}$ tvoria netriviálny rozklad automatu $A_i^D$. Teda $L'_i$ je deterministicky rozložiteľný pre ľubovoľné $i \geq 2$.

Zhrňme, čo sme dokázali. Pre postupnosť jazykov $(L'_i)_{i=2}^{\infty}$ platí, že obsahuje nekonečne veľa jazykov, ktoré sú súčasne nedeterministicky nerozložiteľné a deterministicky rozložiteľné. Sú to tie $L'_i$ pre ktoré je $i$ mocninou prvočísla. Hľadanú postupnosť $(L_i)_{i=2}^{\infty}$ teda získame tak, že z postupnosti $(L'_i)_{i=2}^{\infty}$ vyberieme tie $L'_i$, kde $i$ je mocninou prvočísla. Zjavne postupnosť $(L_i)_{i=2}^{\infty}$ spĺňa (a) a pri lepšom pohľade na dôkaz deterministickej rozložiteľnosti jazykov $L'_i$ zistíme, že spĺňa aj (b).

$\square$

# Kapitola 4

# Vlastnosti rozložiteľnosti a nerozložiteľnosti

Skúmame uzáverové vlastnosti tried rozložiteľných a nerozložiteľných jazykov. Charakterizujeme triedu jednoslovných jazykov vzhľadom na rozložiteľnosť. Skúmame jazyky, ktorých minimálne NKA sú tvorené práve jedných cyklom. Ukazujeme, že ak je minimálny NKA pre jazyk príliš malý, tak je automat automaticky nerozložiteľný. Uvádzame výsledok, ktorý hovorí o jazykoch, ktoré sa navzájom líšia iba v tom, či obsahujú alebo neobsahujú nejaký symbol.

## 4.1 Príliš malé nedeterministické konečné automaty

Dokazujeme, že ak je minimálny NKA pre jazyk príliš malý, tak je jazyk automaticky nerozložiteľný.

**Veta 4.1.1.** *Nech $L$ je jazyk, pričom $nsc(L) \leq 2$. Potom $L$ je nerozložiteľný.*

*Dôkaz.* Pre $nsc(L) = 1$ je tvrdenie zrejmé. Uvažujme $nsc(L) = 2$ a postupujme sporom. Nech je $L$ rozložiteľný, t.j. existujú NKA $A_1$ a $A_2$ také, že $L(A_1) \cap L(A_2) = L$, $\#_S(A_1) = 1$, $\#_S(A_2) = 1$. Pozrime sa však lepšie na to, čo dokážu jednostavové NKA. Dá sa ľahko nahliadnuť, že jednostavový NKA môže akceptovať iba jeden z nasledovných troch typov jazykov: $\emptyset, \{\varepsilon\}, \Sigma^*$, kde $\Sigma$ je ľubovoľná abeceda. Taktiež platí $\emptyset \subset \{\varepsilon\} \subset \Sigma^*$. Z toho vyplýva, že $L(A_1) \cap L(A_2) \in \{\emptyset, \{\varepsilon\}, \Sigma^*\}$. Platí $nsc(\emptyset) = nsc(\{\varepsilon\}) = nsc(\Sigma^*) = 1$, teda $nsc(L(A_1) \cap L(A_2)) = 1$. Avšak $L(A_1) \cap L(A_2) = L$ a podľa predpokladu $nsc(L) = 2$, čo je hľadaný spor. $\square$

## 4.2 Nový symbol v jazyku

Nasledujúca veta formalizuje fakt, že ak máme regulárny jazyk a z neho vytvoríme nový jazyk tak, že vezmeme nový symbol, ktorý slová z pôvodného jazyka neobsahujú a tento symbol „ vopcháme “ do slov pôvodného jazyka, tak na rozložiteľnosti pôvodného jazyka to nič nezmení.

**Veta 4.2.1.** *Nech* $L \in \mathscr{R}$ *a* $b \notin \Sigma_L$*. Definujeme homomorfizmus* $h_b : \Sigma_L \cup \{b\} \to \Sigma_L$ *nasledovne -* $h_b(b) = \varepsilon$, $\forall a \in \Sigma_L : h_b(a) = a$*. Potom platia nasledovné tvrdenia:*

*(a)* $nsc(L) = nsc(h_b^{-1}(L))$

*(b)* $L$ *je rozložiteľný* $\Leftrightarrow$ $h_b^{-1}(L)$ *je rozložiteľný*

*Dôkaz.* Najprv dokážeme (a). Nech $A^L_{min} = (K_L, \Sigma_L, \delta_L, q_L, F_L)$ je minimálny NKA pre $L$. Definujeme NKA $A^b_{min} = (K_L, \Sigma_L \cup \{b\}, \delta_b, q_L, F_L)$ kde $\delta_b$ je definovaná nasledovne - $\forall a \in \Sigma_L\ \forall q \in K_L : \delta_b(q, a) = \delta_L(q, a)$, $\forall q \in K_L : \delta_b(q, b) = \{q\}$. Ako možno ľahko vidieť, do NKA pre $L$ sme iba pridali slučku na $b$ v každom stave a preto platí $L(A^b_{min}) = h_b^{-1}(L)$.

Tvrdíme, že $A^b_{min}$ je minimálny NKA pre $h_b^{-1}(L)$. Toto tvrdenie dokážeme sporom. Nech existuje NKA $A^b_\downarrow = (K^b_\downarrow, \Sigma^b_\downarrow, \delta^b_\downarrow, q^b_\downarrow, F^b_\downarrow)$ taký, že $L(A^b_\downarrow) = h_b^{-1}(L), \#_S(A^b_\downarrow) < \#_S(A^b_{min})$. Na základe $A^b_\downarrow$ definujeme NKA $A^L_\downarrow = (K^b_\downarrow, \Sigma^b_\downarrow - \{b\}, \delta^L_\downarrow, q^b_\downarrow, F^b_\downarrow)$ kde prechodová funkcia $\delta^L_\downarrow$ je definovaná nasledovne - $\forall q \in K^b_\downarrow\ \forall a \in \Sigma^b_\downarrow - \{b\} : \delta^L_\downarrow(q, a) = \delta^b_\downarrow(q, a)$. Dokážeme, že $L(A^L_\downarrow) = L$.
$\subseteq$: Nech $w \in L(A^L_\downarrow)$. Potom existuje akceptačný výpočet na $w$ v automate $A^L_\downarrow$. Vďaka tomu, ako je $A^L_\downarrow$ definovaný je tento výpočet taktiež akceptačným výpočtom v automate $A^b_\downarrow$ a teda $w \in h_b^{-1}(L)$, z čoho plynie $h_b(w) \in L$. Avšak z toho ako je $A^L_\downarrow$ definovaný vyplýva, že $w$ neobsahuje symbol $b$ a teda $h_b(w) = w$ z čoho plynie $w \in L$
$\supseteq$: Nech $w \in L$. Z toho ľahko vidno, že $w \in h_b^{-1}(L)$. Teda existuje akceptačný výpočet na slove $w$ v automate $A^b_\downarrow$. Nakoľko $w$ neobsahuje symbol $b$ a automat $A^L_\downarrow$ obsahuje všetky prechody automatu $A^b_\downarrow$ okrem prechodov na $b$, tak zmienený výpočet je taktiež akceptačným výpočtom na slove $w$ v automate $A^L_\downarrow$, čo kompletizuje dôkaz tvrdenia $L(A^L_\downarrow) = L$.
Z predošlého vyplýva $\#_S(A^L_\downarrow) = \#_S(A^b_\downarrow) < \#_S(A^b_{min}) = \#_S(A^L_{min})$, čo je v spore s predpokladom, že automat $A^L_{min}$ je minimálny NKA pre jazyk $L$. Teda automat $A^b_\downarrow$ s uvedenými vlastnosťami nemôže existovať a teda $A^b_{min}$ je minimálny NKA pre $h_b^{-1}(L)$. Z konštrukcie automatu $A^b_{min}$ plynie, že $\#_S(A^b_{min}) = \#_S(A^L_{min})$, čo kompletizuje dôkaz (a).

Dokážeme tvrdenie (b).
$\Rightarrow$: Nech je $L$ rozložiteľný. Teda ak $A^L_{min}$ je minimálny NKA pre $L$, tak existuje jeho netriviálny rozklad na NKA $A^L_1 = (K_1, \Sigma_1, \delta_1, q_1, F_1)$ a $A^L_2 = (K_2, \Sigma_2, \delta_2, q_2, F_2)$. BUNV môžeme predpokladať, že $b \notin \Sigma_1, b \notin \Sigma_2$. Definujeme NKA $A^b_1 = (K_1, \Sigma_1 \cup \{b\}, \delta^b_1, q_1, F_1)$ kde prechodová funkcia $\delta^b_1$ je definovaná nasledovne - $\forall q \in K_1\ \forall a \in \Sigma_1 : \delta^b_1(q, a) = \delta_1(q, a)$, $\forall q \in K_1 : \delta^b_1(q, b) = \{q\}$. Ako si možno všimnúť, automat $A^b_1$ sme zostrojili z automatu $A^L_1$ tak, že sme v každom stave pridali slučku na

$b$ a teda ľahko vidno, že $L(A_1^b) = h_b^{-1}(L(A_1^L))$. Analogicky vieme definovať na základe $A_2^L$ NKA $A_2^b$ o ktorom analogicky platí $L(A_2^b) = h_b^{-1}(L(A_2^L))$. Označme minimálny NKA pre jazyk $h_b^{-1}(L)$ $A_{min}^b$. Podľa (a) platí $\#_S(A_{min}^b) = \#_S(A_{min}^L)$. Nakoľko $\#_S(A_1^L) = \#_S(A_1^b)$ a $\#_S(A_2^L) = \#_S(A_2^b)$, tak na to, aby sme dokázali, že $A_1^b$ a $A_2^b$ tvoria netriviálny rozklad automatu $A_{min}^b$ stačí dokázať $L(A_1^b) \cap L(A_2^b) = h_b^{-1}(L)$. To dokážeme nasledujúcou argumentáciou, ktorá vyplýva z vlastností inverzných homomorfizmov a konštrukcie automatov, ktoré v dôkaze používame - $w \in L(A_1^b) \cap L(A_2^b) \Leftrightarrow w \in h_b^{-1}(L(A_1^L)) \cap h_b^{-1}(L(A_2^L)) \Leftrightarrow w \in h_b^{-1}(L(A_1^L) \cap L(A_2^L)) \Leftrightarrow w \in h_b^{-1}(L)$. Teda $h_b^{-1}(L)$ je rozložiteľný.

$\Leftarrow$: Nech $h_b^{-1}(L)$ je rozložiteľný. Nech $A_{min}^b$ je minimálny NKA pre $h_b^{-1}(L)$. Teda existuje netriviálny rozklad automatu $A_{min}^b$. Nech NKA tvoriace tento rozklad sú $A_1^b = (K_1^b, \Sigma_1^b, \delta_1^b, q_1^b, F_1^b)$ a $A_2^b = (K_2^b, \Sigma_2^b, \delta_2^b, q_2^b, F_2^b)$. Nech $A_{min}^L$ je minimálny NKA pre jazyk $L$. Chceme skonštruovať netriviálny rozklad automatu $A_{min}^L$. Na základe $A_1^b$ definujeme NKA $A_1^L = (K_1^b, \Sigma_1^b - \{b\}, \delta_1^L, q_1^b, F_1^b)$ kde prechodová funkcia $\delta_1^L$ je definovaná nasledovne - $\forall q \in K_1^b \forall a \in \Sigma_1^b - \{b\} : \delta_1^L(q,a) = \delta_1^b(q,a)$. Hlavnou myšlienkou je, že z automatu $A_1^b$ sme vynechali prechody na $b$, pretože ich v rozklade, ktorý chceme vytvoriť, aj tak nepotrebujeme. Analogicky, na základe $A_2^b$, definujeme NKA $A_2^L$. Dokážeme, že $L = L(A_1^L) \cap L(A_2^L)$.

$\subseteq$: Nech $w \in L$. Potom aj $w \in h_b^{-1}(L)$. Teda $w \in L(A_1^b) \cap L(A_2^b)$. Nakoľko však $w$ neobsahuje symbol $b$, tak z konštrukcie $A_1^L$ plynie $w \in L(A_1^L)$ (pretože $A_1^L$ obsahuje všetky prechody z $A_1^b$ okrem prechodov na $b$, ktoré ale pri výpočte na $w$ nepotrebujeme). Analogicky $w \in L(A_2^L)$. Teda $w \in L(A_1^L) \cap L(A_2^L)$.

$\supseteq$: Nech $w \in L(A_1^L) \cap L(A_2^L)$. Nakoľko automat $A_1^b$ respektíve $A_2^b$ obsahuje všetky prechody, ktoré obsahuje automat $A_1^L$ respektíve $A_2^L$, tak $w \in L(A_1^b) \cap L(A_2^b)$. Teda $w \in h_b^{-1}(L)$, teda $h_b(w) \in L$. Nakoľko $w$ neobsahuje symbol $b$, tak $h_b(w) = w$, z čoho plynie $w \in L$. Takže $L = L(A_1^L) \cap L(A_2^L)$.

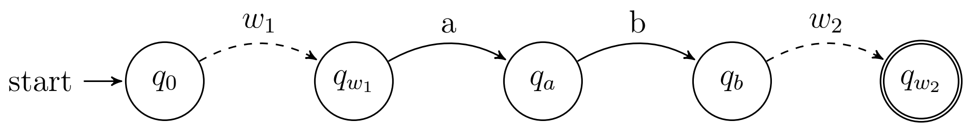
Kedže $L = L(A_{min}^L)$, tak platí $L(A_{min}^L) = L(A_1^L) \cap L(A_2^L)$. Z (a) vyplýva $\#_S(A_{min}^L) = \#_S(A_{min}^b)$. Z konštrukcie $A_1^L$ respektíve $A_2^L$ vyplýva $\#_S(A_1^L) = \#_S(A_1^b)$ respektíve $\#_S(A_2^L) = \#_S(A_2^b)$. Nakoľko $A_1^b$ a $A_2^b$ tvoria netriviálny rozklad automatu $A_{min}^b$, tak platí $\#_S(A_1^b) < \#_S(A_{min}^b)$ a $\#_S(A_2^b) < \#_S(A_{min}^b)$. Z toho vyplýva $\#_S(A_1^L) < \#_S(A_{min}^L)$ a $\#_S(A_2^L) < \#_S(A_{min}^L)$. Teda automaty $A_1^L$ a $A_2^L$ tvoria netriviálny rozklad automatu $A_{min}^L$. Teda jazyk $L$ je rozložiteľný. □

## 4.3 Charakterizácia jazykov tvorených jedným slovom

Uvádzame úplnú charakterizáciu triedy jazykov tvorených práve jedným slovom vzhľadom na rozložiteľnosť.

**Veta 4.3.1.** *Nech $L = \{w\}$. Potom je $L$ rozložiteľný práve vtedy, keď $w$ obsahuje aspoň dva rôzne symboly.*

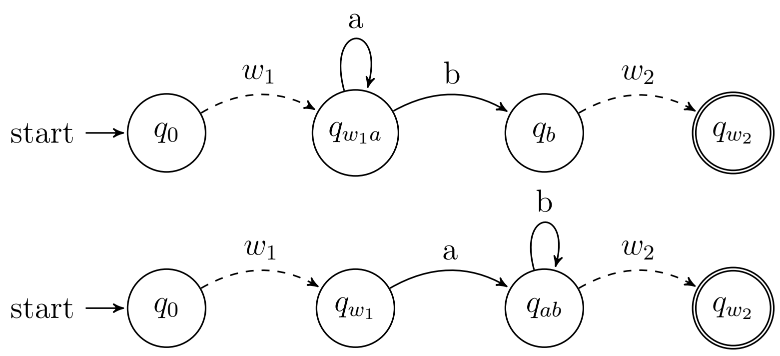
*Dôkaz.* $\Rightarrow$: Dokážeme obmenu tvrdenia. Ak $w$ obsahuje nanajvýš jeden znak, tak existuje nejaké $n \in \mathbb{N}$ také, že $L = \{a^n\}$. Potom podľa Tvrdenia 2.2.1 je jazyk $L$ nerozložiteľný. $\Leftarrow$: Nech $w = w_1abw_2$ pre nejaké slová $w_1, w_2$. Zostrojíme NKA $A_w$ pre jazyk $L = \{w\}$. Automat uvádzame pomocou zovšeobecneného diagramu.




Obr. 4.1: automat $A_w$

Uvažujme množinu dvojíc slov $F = \{(pref(w,i), suff(w,|w|-i)) \mid 0 \leq i \leq |u|\}$. Množina $F$ je podľa definície 1.4.1 mätúcou množinou pre jazyk $L$. Nakoľko $|F| = |w|+1$, tak podľa Vety 1.4.1 $nsc(L) \geq |w|+1$. Kedže $L(A_w) = L$ a $\#_S(A_w) = |w|+1$, tak $nsc(L) = |w|+1$ a automat $A_w$ je minimálny NKA pre jazyk $L$.

Zostrojíme netriviálny rozklad automatu $A_w$. Hľadané automaty $A_w^a$ a $A_w^b$ uvádzame pomocou zovšeobecnených diagramov.




Obr. 4.2: Rozklad automatu $A_w$ na automaty $A_w^a$ (hore) a $A_w^b$ (dole)

Ľahko vidno, že $L(A_w^a) = \{w_1a^kbw_2 | k \in \mathbb{N}\}$ a $L(A_w^b) = \{w_1ab^kw_2 | k \in \mathbb{N}\}$. Ukážeme, že $L(A_w^a) \cap L(A_w^b) = \{w\}$.
$\subseteq$: Nech $u \in L(A_w^a) \cap L(A_w^b)$. Teda existujú $k_1, k_2 \in \mathbb{N}$ také, že $u = w_1a^{k_1}bw_2 = w_1ab^{k_2}w_2$. Musí platiť $k_1 = k_2$, inak by platilo $|u| \neq |u|$. Taktiež musí platiť $k_1 \geq 1$, lebo $pref(u, |w_1|+1) = w_1a$. Teda aj $k_2 \geq 1$, lebo inak by bolo $k_1 \neq k_2$. Teda $pref(u, |w_1|+2) = w_1ab$. Z toho nutne $k_1 = 1$ a teda aj $k_2 = 1$. Takže $u = w_1abw_2 = w$.
$\supseteq$: Táto inklúzia je zjavná.

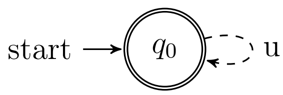
Navyše platí $\#_S(A_w^a) < \#_S(A_w)$ a $\#_S(A_w^b) < \#_S(A_w)$. Takže $A_w^a$ a $A_w^b$ tvoria netriviálny rozklad automatu $A_w$, čo ukazuje, že jazyk $L$ je rozložiteľný. $\square$

## 4.4 Automaty tvorené jediným cyklom

Typickou schopnosťou konečných automatov je počítať v cykle zvyšok po delení daným číslom z dĺžky slova. Tieto automaty sa vyznačujú tým, že sú tvorené jediným cyklom, pričom nijak nezohľadňujú štruktúru slova. Podstatu otázok spojených s takýmito automatmi riešia Vety 2.2.1 a 2.1.1. Nakoľko v konečných automatoch sú práve cykly veľmi dôležitou štruktúrou, v našej práci sme túto otázku rozšírili a študovali sme otázku rozložiteľnosti jazykov, ktorých minimálne nedeterministické konečné automaty sú tvorené jediným cyklom, pričom v ňom zohľadňujú aj štruktrúru akceptovaného slova. Podstatou týchto automatov je, neformálne povedané, pumpovanie nejakého slova.

**Lema 4.4.1.** *Nech $\Sigma$ je ľubovolná abeceda, nech $u \in \Sigma^*$, nech $L_u = \{u\}^*$. Potom $nsc(L_u) = |u|$.*

*Dôkaz.* Zostrojíme NKA $A_u$ pre jazyk $L_u$. Tento automat s počtom stavov $|u|$ uvádzame pomocou zovšeobecneného diagramu.

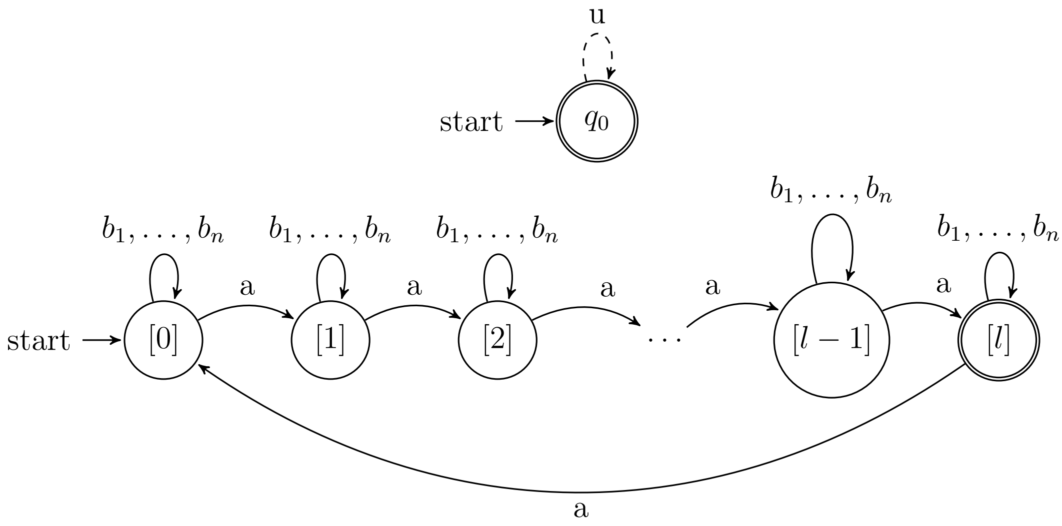



Obr. 4.3: automat $A_u$

Ľahko vidno, že $L(A_u) = L_u$. Uvažujme množinu dvojíc slov $F = \{(pref(u,i),\ suff(|u|-i)) \mid 0 \leq i < |u|\}$. Množina $F$ je podľa definície 1.4.1 mätúcou množinou pre jazyk $L_u$. Nakoľko $|F| = |u|$, tak podľa Vety 1.4.1 $nsc(L_u) \geq |u|$. Kedže $L(A_u) = L_u$ a $\#_S(A_u) = |u|$, tak $nsc(L_u) = |u|$ a automat $A_u$ je minimálny NKA pre jazyk $L_u$. □

**Tvrdenie 4.4.1.** *Nech $\Sigma$ je ľubovolná abeceda taká, že $|\Sigma| \geq 2$. Nech pre $u \in \Sigma^*$, $k \geq 2$ je $L_u^k = \{u^k\}^*$. Ak $u$ obsahuje aspoň dva rôzne symboly, potom je $L_u^k$ rozložiteľný.*

*Dôkaz.* Nech $n \geq 1, \Sigma = \{a, b_1, \ldots, b_n\}, u \in \Sigma^*$, $u$ obsahuje symbol $a$ a minimálne ešte jeden symbol zo $\Sigma$. Podľa Lemy 4.4.1 platí $nsc(L_u^k) = k|u|$. Teda existuje NKA $A_u^k$ taký, že $L(A_u^k) = L_u^k$ a $\#_S(A_u^k) = k|u|$. Automat $A_u^k$ je teda minimálny NKA pre $L_u^k$. Zostrojíme netriviálny rozklad automatu $A_u^k$. Označme $l = k.\#_a(u)$. Rozklad uvádzame pomocou zovšeobecneného diagramu.




Obr. 4.4: rozklad automatu $A_u^k$ na automaty $A_u$(hore) a $A_k$(dole)

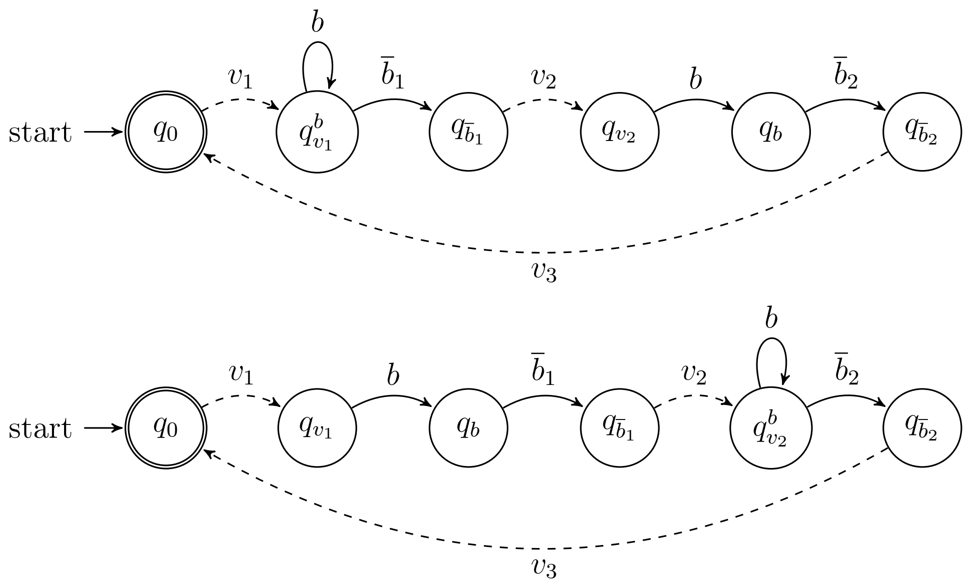
Myšlienkou tohto rozkladu je, že jeden z automatov kontroluje štruktúru slova, či je práve niekoľkonásobným zreťazením slova $u$ a druhý automat kontroluje, či je slov $u$ správne veľa. To však robí tak, že iba počíta počet nejakého jedného symbolu (v našom prípade ho označujeme $a$), ktorý $u$ obsahuje, pričom kontroluje, či slovo obsahuje práve $m.k.\#_a(u)$ pre nejaké $m \in \mathbb{N}$. Formálne $L(A_u) = \{u\}^*$ a $L(A_k) = \{w \in \Sigma^* \mid \#_a(w) \equiv 0 \ (mod\ k.\#_a(u))\}$. Teda $L(A_u) \cap L(A_k) = L(A_u^k)$. Navyše $\#_S(A_u) < \#_S(A_u^k)$ a $\#_S(A_k) < \#_S(A_u^k)$. Je dobré si uvedomiť, že kvôli prvej nerovnosti potrebujeme predpoklad $k \geq 2$ a kvôli druhej nerovnosti potrebujeme predpoklad o veľkosti abecedy $\Sigma$. Teda automaty $A_u$ a $A_k$ tvoria netriviálny rozklad automatu $A_u^k$.

□

**Tvrdenie 4.4.2.** *Nech $\Sigma$ je ľubovolná abeceda, nech $k_1, k_2 \in \{0, 1\}$, nech $w_1, w_2, w_3, w_4, w_5, w_6 \in \Sigma^*$. Definujeme $L = \{w_1 a^{k_1} w_2 b w_3 a w_4 b w_5 a^{k_2} w_6\}^*$. Ak $k_1 = 1$ alebo $k_2 = 1$, potom je $L$ rozložiteľný.*

*Dôkaz.* Zaveďme označenia $u = w_1 a^{k_1} w_2 b w_3 a w_4 b w_5 a^{k_2} w_6$ a $\Sigma_{ab} = \Sigma \cup \{a, b\}$. Podľa Lemy 4.4.1 platí $nsc(L) = |u|$. Teda existuje NKA $A$ taký, že $L(A) = L$ s $\#_S(A) = |u|$. Automat $A$ je teda minimálny NKA pre $L$. Zostrojíme netriviálny rozklad automatu $A$. Rozoberieme nasledujúce dva prípady, podľa toho akého tvaru je slovo $u$. Podľa predpokladov je $u$ práve jedého z nasledujúcich tvarov:

1. Existujú dve rôzne podslová v slove $u$ také, že symbol $b$ je nasledovaný symbolom rôznym od $b$. Formálne existujú $v_1, v_2, v_3 \in \Sigma_{ab}^*$ a $\bar{b}_1, \bar{b}_2 \in \Sigma_{ab} - \{b\}$ také, že $u = v_1 b \bar{b}_1 v_2 \bar{b}_2 v_3$. Na základe tohto poznatku zostrojíme netriviálny rozklad automatu $A$. Rozklad uvádzame pomocou zovšeobecneného diagramu.




Obr. 4.5: rozklad automatu $A$ na automaty $A_1$ a $A_2$

Možno nahliadnuť, že $L(A_1) = \{v_1 b^l \overline{b}_1 v_2 b \overline{b}_2 v_3 \mid l \in \mathbb{N}\}^*$ a $L(A_2) = \{v_1 b \overline{b}_1 v_2 b^l \overline{b}_2 v_3 \mid l \in \mathbb{N}\}^*$.

Dokážeme, že $L(A_1) \cap L(A_2) = L$.
$\supseteq$: Táto inklúzia je triviálna, nebudeme ju formálne dokazovať.
$\subseteq$: Uvažujme $w \in L(A_1) \cap L(A_2)$. Potom existuje $n, m, l_1, \ldots, l_n, o_1, \ldots, o_m \in \mathbb{N}$ také, že $w = v_1 b^{l_1} \overline{b}_1 v_2 b \overline{b}_2 v_3 \ldots v_1 b^{l_n} \overline{b}_1 v_2 b \overline{b}_2 v_3 = v_1 b \overline{b}_1 v_2 b^{o_1} \overline{b}_2 v_3 \ldots v_1 b \overline{b}_1 v_2 b^{o_m} \overline{b}_2 v_3$. Indukciou na $n$ dokážeme, že $m = n$, pre $0 \leq i \leq n : l_i = 1$ a pre $0 \leq i \leq m : o_i = 1$.

$1^0$ : Ak $n = 0$, tak $w = \varepsilon$ a tvrdenie triviálne platí.

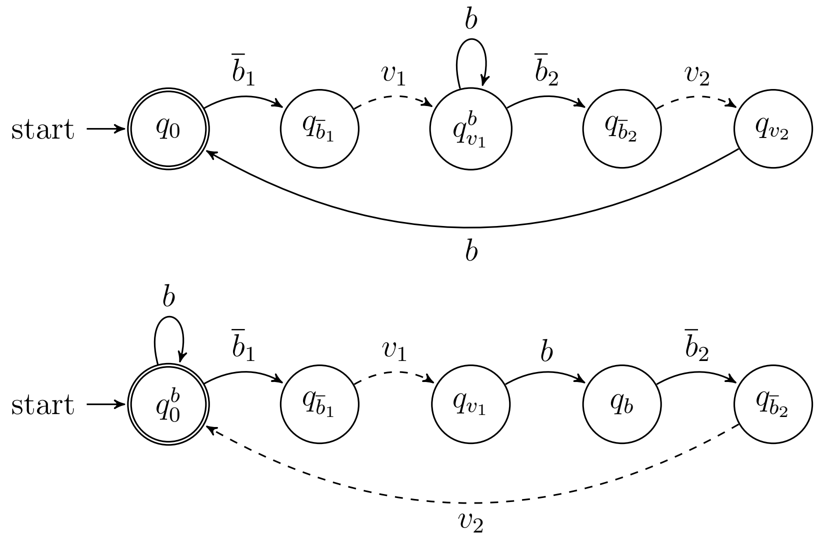
$2^0$ : Platí $v_1 b^{l_1} \overline{b}_1 v_2 b \overline{b}_2 v_3 \ldots v_1 b^{l_n} \overline{b}_1 v_2 b \overline{b}_2 v_3 = v_1 b \overline{b}_1 v_2 b^{o_1} \overline{b}_2 v_3 \ldots v_1 b \overline{b}_1 v_2 b^{o_m} \overline{b}_2 v_3$. Pozrime sa pozornejšie na prvé úseky v tomto slove, t.j. na časti $v_1 b^{l_1} \overline{b}_1 v_2 b \overline{b}_2 v_3$ a $v_1 b \overline{b}_1 v_2 b^{o_1} \overline{b}_2 v_3$. Oba úseky sú prefixom toho istého slova a na prvých $|v_1|$ symboloch sa evidentne zhodujú. Musí platiť $l_1 \geq 1$, aby sa zhodovali aj na symoble $b$, ktorý nasleduje za $v_1$. Avšak nakoľko v tomto prefixe po zmienenom $b$ nasleduje znak $\overline{b}_1$, tak nutne $l_1 = 1$. Teda platí $v_1 b^{l_1} \overline{b}_1 v_2 = v_1 b \overline{b}_1 v_2$. Z toho plynie $o_1 \geq 1$, nakoľko po $v_2$ musí nasledovať symbol $b$. Ďalším sybmobolom je však $\overline{b}_2$, teda nutne $o_1 = 1$. Teda platí $v_1 b^{l_1} \overline{b}_1 v_2 b \overline{b}_2 v_3 = v_1 b \overline{b}_1 v_2 b^{o_1} \overline{b}_2 v_3 = v_1 b \overline{b}_1 v_2 b \overline{b}_2 v_3$. Navyše, oba automaty, $A_1$ aj $A_2$ sa po dočítaní tohto prefixu dostanú práve do ich počiatočného (a zároveň jediného akceptačného) stavu $q_0$. V prípade, že $n = 1$, tak niet čo ďalej dokazovať. Ak $n \geq 2$ tak z predchádzajúceho vyplýva, že $v_1 b^{l_2} \overline{b}_1 v_2 b \overline{b}_2 v_3 \ldots v_1 b^{l_n} \overline{b}_1 v_2 b \overline{b}_2 v_3 = v_1 b \overline{b}_1 v_2 b^{o_2} \overline{b}_2 v_3 \ldots v_1 b \overline{b}_1 v_2 b^{o_m} \overline{b}_2 v_3$ a navyše toto slovo akceptujú oba automaty,

$A_1$ aj $A_2$. Teda podľa indukčného predpokladu môžeme tvrdiť, že $n = m$, pre $2 \leq i \leq n$ platí $l_i = o_i = 1$, čo dokazuje tvrdenie.

Z predošlého vyplýva, že $w \in L$, čo kompletizuje dôkaz tejto inklúzie.

Teda $L(A_1) \cap L(A_2) = L = L(A)$. Navyše $\#_S(A_1) < \#_S(A)$ a $\#_S(A_2) < \#_S(A)$, teda automaty $A_1$ a $A_2$ tvoria netriviálny rozklad automatu $A$.

2. Existujú $\overline{b}_1, \overline{b}_2 \in \Sigma_{ab} - \{b\}$, $v_1, v_2 \in \Sigma_{ab} - \{b\}$, $c_1, c_2 \geq 1$ také, že $u = \overline{b}_1 v_1 b^{c_1} \overline{b}_2 v_2 b^{c_2}$. Na základe tohto poznatku zostrojíme netriviálny rozklad automatu $A$. Rozklad uvádzame pomocou zovšeobecneného diagramu.




Obr. 4.6: rozklad automatu $A$ na automaty $A_1$(hore) a $A_2$(dole)

Možno nahliadnuť, že $L(A_1) = \{\overline{b}_1 v_1 b^l \overline{b}_2 v_2 b \mid l \in \mathbb{N}\}^*$ a $L(A_2) = \{b^l \overline{b}_1 v_1 b \overline{b}_2 v_2 \mid l \in \mathbb{N}\}^* \{b\}^*$. Platí $L(A_1) \cap L(A_2) = L$, čo sa dá dokázať veľmi podobne a rovnako veľmi technicky ako v predošlom prípade, preto dôkaz neuvádzame. Navyše $\#_S(A_1) < \#_S(A)$ a $\#_S(A_2) < \#_S(A)$, teda automaty $A_1$ a $A_2$ tvoria netriviálny rozklad automatu $A$.

Záverom ešte spomeňme, že hlavnou myšlienkou rozkladu bola akási synchronizácia výpočtov automatov v rozklade na symboloch rôznych od $b$, ktoré nasledovali hneď za $b$. $\square$

## 4.5 Uzáverové vlastnosti

Skúmame uzáverové vlastnosti tried rozložiteľných a nerozložiteľných jazykov. Ukazujeme, že obe triedy nie sú uzavreté na žiadnu zo štandardných operácií.

**Tvrdenie 4.5.1.** *Trieda rozložiteľných jazykov nie je uzavretá na prienik.*

*Dôkaz.* Uvažujme jazyky $L_1 = \{a^{92}\} \cup \{b\}^*, L_2 = \{a^{92}\} \cup \{c\}^*$. $L_1$ a $L_2$ sú podla Tvrdenia 2.1.2 rozložiteľné. Avšak jazyk $L_1 \cap L_2 = \{a^{92}\}$ je podľa Tvrdenia 2.2.1 nerozložiteľný. □

**Tvrdenie 4.5.2.** *Trieda nerozložiteľných jazykov nie je uzavretá na prienik.*

*Dôkaz.* Uvažujme jazyky $L_1 = \{a^{2017k} \mid k \in \mathbb{N}\}, L_2 = \{a^{29k} \mid k \in \mathbb{N}\}$. $L_1$ a $L_2$ sú podla Vety 2.2.1 nerozložiteľné. Avšak jazyk $L_1 \cap L_2 = \{a^{58493k} \mid k \in \mathbb{N}\}$ je podľa Vety 2.1.1 rozložiteľný. □

**Tvrdenie 4.5.3.** *Trieda rozložiteľných jazykov nie je uzavretá na zjednotenie.*

*Dôkaz.* Uvažujme jazyky $L_1 = \{w \in \{a,b\}^* \mid \#_a(w) \equiv 0\ (mod\ 2),\ \#_b(w) \equiv 0\ (mod\ 3)\}, L_2 = \{w \in \{a,b\}^* \mid \#_a(w) \equiv 1\ (mod\ 2),\ \#_b(w) \equiv 0\ (mod\ 3)\}$. $L_1$ a $L_2$ sú podľa Tvrdenia 2.1.4 rozložiteľné. Avšak jazyk $L_1 \cup L_2 = \{w \in \{a,b\}^* \mid \#_b(w) \equiv 0\ (mod\ 3)\}$ je podľa Vety 2.2.1 a Vety 4.2.1 nerozložiteľný. □

**Tvrdenie 4.5.4.** *Trieda nerozložiteľných jazykov nie je uzavretá na zjednotenie.*

*Dôkaz.* Uvažujme jazyky $L_1 = \{a^{2829}\}, L_2 = \{b\}^*$. Podľa Tvrdenia 2.2.1 je $L_1$ je nerozložiteľný a $L_2$ je podla Vety 4.1.1 nerozložiteľný. Avšak jazyk $L_1 \cup L_2$ je podľa Tvrdenia 2.1.2 rozložiteľný. □

**Tvrdenie 4.5.5.** *Trieda rozložiteľných jazykov nie je uzavretá na homomorfizmus.*

*Dôkaz.* Uvažujme jazyk $L = \{a^{89}\} \cup \{b\}^*$ a homomorfizmus $h : \{a,b\} \to \{a\}$ definovaný nasledovne - $h(a) = a, h(b) = a$. Jazyk $L$ je podľa Tvrdenia 2.1.2 rozložiteľný. Avšak jazyk $h(L) = \{a\}^*$ je podľa Vety 4.1.1 nerozložiteľný. □

**Tvrdenie 4.5.6.** *Trieda nerozložiteľných jazykov nie je uzavretá na homomorfizmus.*

*Dôkaz.* Uvažujme jazyk $L = \{a^{2k} | k \in \mathbb{N}\}$ a homomorfizmus $h : \{a\} \to \{a\}$ definovaný nasledovne - $h(a) = aaa$. Jazyk $L$ je podľa Vety 2.2.1 nerozložiteľný. Avšak jazyk $h(L) = \{a^{6k} | k \in \mathbb{N}\}$ je podľa Vety 2.1.1 rozložiteľný. □

**Tvrdenie 4.5.7.** *Trieda rozložiteľných jazykov nie je uzavretá na inverzný homomorfizmus.*

*Dôkaz.* Uvažujme jazyk $L = \{a^{39}\} \cup \{b\}^*$ a homomorfizmus $h : \{b\} \to \{b\}$ definovaný nasledovne - $h(b) = b$. Jazyk $L$ je podľa Tvrdenia 2.1.2 rozložiteľný. Avšak jazyk $h^{-1}(L) = \{b\}^*$ je podľa Vety 4.1.1 nerozložiteľný. □

**Tvrdenie 4.5.8.** *Trieda rozložiteľných jazykov nie je uzavretá na zreťazenie.*

*Dôkaz.* Uvažujme jazyky $L_1 = \{w \in \{a,b\}^* \mid \#_a(w) \equiv 0 \ (mod\ 3),\ \#_b(w) \equiv 0 \ (mod\ 2)\}, L_2 = \{w \in \{a,b\}^* \mid \#_a(w) \equiv 0 \ (mod\ 3),\ \#_b(w) \equiv 1 \ (mod\ 2)\} \cup \{\varepsilon\}$. Jazyky $L_1$ a $L_2$ sú podľa Tvrdenia 2.1.4 rozložiteľné. Platí $L_1.L_2 = \{w \in \{a,b\}^* \mid \#_a(w) \equiv 0 \ (mod\ 3)\}$. Teda jazyk $L_1.L_2$ je podľa Vety 2.2.1 a Vety 4.2.1 nerozložiteľný. □

**Tvrdenie 4.5.9.** *Trieda nerozložiteľných jazykov nie je uzavretá na zreťazenie.*

*Dôkaz.* Uvažujme jazyky $L_1 = \{b\}, L_2 = \{w \in \{a,b\}^* \mid \#_a(w) = 81\}$. $L_1$ je podľa Vety 4.1.1 nerozložiteľný a $L_2$ je v podľa Tvrdenia 2.2.1 a Vety 4.2.1 nerozložiteľný. Avšak jazyk $L_1.L_2$ je podľa Tvrdenia 2.1.3 rozložiteľný. □

**Tvrdenie 4.5.10.** *Trieda rozložiteľných jazykov nie je uzavretá na iteráciu.*

*Dôkaz.* Uvažujme jazyk $L = \{ab\}$. Jazyk $L$ je podľa 4.3.1 rozložiteľný. Podľa lemy 4.4.1 platí $nsc(L^*) = 2$ a teda podľa Vety 4.1.1 je jazyk $L^*$ nerozložiteľný. □

**Tvrdenie 4.5.11.** *Trieda nerozložiteľných jazykov nie je uzavretá na iteráciu.*

*Dôkaz.* Uvažujme jazyk $L = \{a^{15}\}$. Jazyk $L$ je podľa 4.3.1 nerozložiteľný. $L^* = \{a^{15k} \mid k \in \mathbb{N}\}$, teda $L^*$ je podľa Vety 2.1.1 rozložiteľný. □

Zostáva otvorené, či je trieda nerozložiteľných jazykov uzavretá na inverzný homomorfizmus.

# Záver

V práci sme nadviazali na výskum v oblasti skúmania rôznych aspektov informácie. Skúmali sme pojem užitočnosti informácie. Otvorili sme oblasť skúmania užitočnosti prídavnej informácie v kontexte nedeterminizmu. Ako výpočtový model sme zvolili nedeterministické konečné automaty s mierou zložitosti počet stavov. Formalizáciou nášho problému je rozklad nedeterministického konečného automatu. Pojem rozložiteľnosti sme prirodzene rozšírili na regulárne jazyky.

V práci sme dokázali rozložiteľnosť resp. nerozložiteľnosť niekoľkých konkrétnych regulárnych jazykov. Tieto výsledky pomáhajú uchopiť problém rozložiteľnosti a nerozložiteľnosti a pomáhajú vybudovať dôkazové techniky. Tieto výsledky sme následne použili pri skúmaní uzáverových vlastností rozložiteľných resp. nerozložiteľných regulárnych jazykov. Dokázali sme, že tieto triedy nie sú uzavreté na žiadnu z bežných operácií. Charakterizovali sme dve podtriedy regulárnych jazykov vzhľadom na rozložiteľnosť. Sú to jazyky pozostávajúce z jedného slova a jazyky tvaru $\{a^{kn} \mid k \in \mathbb{N}\}$. Ukázali sme rozdiel medzi deterministickou a nedeterministickou rozložiteľnosťou. Našli sme nekonečnú postupnosť regulárnych jazykov, ktoré sú nedeterministicky nerozložiteľné ale deterministicky rozložiteľné. Navyše rozklad minimálneho deterministického konečného automatu pre tieto jazyky je taký, že oba deterministické automaty v rozklade majú asi polovicu stavov vzhľadom k pôvodnému automatu.

Možným pokračovaním tejto práce je hľadanie charakterizácií ďalších netriviálnych podtried regulárnych jazykov vzhľadom na nedeterministickú rozložiteľnosť. Veľmi dobrým výsledkom by bolo nájsť charakterizáciu regulárnych jazykov vzhľadom na rozložiteľnosť. Z uzáverových vlastností by sa dala skúmať ešte uzavretosť na reverz a komplement. Zmysluplným pokračovaním je tiež skúmanie rozložiteľnosti nedeterministického konečného automatu ako takého (neuvažovať v kontexte rozložiteľnosti jazyka) tak, že sa pozrieme na jeho definíciu a z nej skúsime usúdiť, či je daný automat rozložiteľný (dalo by sa pozrieť napr. na grafové vlastnosti diagramu daného automatu). Ďalšou možnosťou je skúmanie pojmu rozložiteľnosti pre iné, silnejšie, výpočtové modely.

# Literatúra

[Gaži and Rovan, 2008] Gaži, P. and Rovan, B. (2008). *Assisted Problem Solving and Decompositions of Finite Automata*, pages 292–303. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Glaister and Shallit, 1996] Glaister, I. and Shallit, J. (1996). A lower bound technique for the size of nondeterministic finite automata. *Information Processing Letters*, (59):75–77.

[Gruber and Holzer, 2006] Gruber, H. and Holzer, M. (2006). Finding lower bounds for nondeterministic state complexity is hard. Technical report, Institut fur Informatik, Technische Universitat Munchen, Boltzmannstraße 3, D-85748 Garching bei Munchen, Germany.

[Labath and Rovan, 2011] Labath, P. and Rovan, B. (2011). *Simplifying DPDA Using Supplementary Information*, pages 342–353. Springer Berlin Heidelberg, Berlin, Heidelberg.

[McCulloch and Pitts, 1943] McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, (5):115–133.

[Palioudakis, 2012] Palioudakis, A. (October 2012). Nondeterministic state complexity and quantifying non-determinism in finite automata. Technical Report Technical Report 2012-596, School of Computing, Queen's University, Kingston, ON, Canada.