Pologrupy a formální jazyky

Michal Kunc

23. září 2024

Obsah

Úvod **1**

1 Rozpoznatelné a racionální podmnožiny pologrup **3**
1.1 Pologrupy a monoidy . . . 3
1.1.1 Volné monoidy . . . 4
1.1.2 Pologrupy transformací . . . 5
1.2 Rozpoznatelné podmnožiny . . . 5
1.2.1 Rozpoznávání homomorfismy . . . 5
1.2.2 Deterministické automaty . . . 7
1.2.3 Rozpoznávání uspořádanými pologrupami . . . 9
1.2.4 Minimální automat a syntaktický homomorfismus . . . 11
1.3 Racionální podmnožiny . . . 14
1.4 Vztahy mezi rozpoznatelnými a racionálními podmnožinami . . . 16
1.5 Uzávěrové vlastnosti . . . 16
1.6 Speciální případy . . . 18
1.6.1 Monoidy slov . . . 19
1.6.2 Rozpoznatelné a racionální relace . . . 20
1.6.3 Grupy . . . 22
1.6.4 Volné komutativní monoidy . . . 23
1.6.5 Monoidy stop* . . . 23

2 Struktura konečných pologrup **25**
2.1 Podgrupy . . . 25
2.2 Jednogenerované podpologrupy . . . 25
2.3 Greenovy relace . . . 26
2.3.1 Definice a základní vlastnosti . . . 27
2.3.2 Greenovy relace v konečných pologrupách . . . 32
2.3.3 Regulární $\mathscr{D}$-třídy . . . 36
2.4 Hlavní faktory konečných pologrup . . . 41
2.4.1 Konečné 0-jednoduché pologrupy . . . 43
2.5 Příklady aplikací . . . 45

2.5.1 Star-free jazyky a aperiodické pologrupy 45
2.5.2 Opakování prvků v součinech a vlastnost konečné mocniny* . . 49
2.6 Faktorizační lesy 50
2.7 Věnčité a kaskádové součiny 57

3 Variety jazyků **65**
3.1 Pseudovariety 65
3.2 Pseudoidentity 68
3.3 Eilenbergova korespondence 75
3.4 Příklady 77

4 Dobrá předuspořádání* **85**
4.1 Definice dobrých předuspořádání 85
4.2 Rozpoznávání jazyků dobrými předuspořádáními 87
4.3 Relace odvozování pro bezkontextové přepisovací systémy 89
4.3.1 Unitární bezkontextové systémy 89
4.3.2 Obecné bezkontextové systémy 90
4.3.3 Bezkontextové systémy definované homomorfismy 90
4.4 Aplikace 92

Literatura **95**

Úvod

Cílem tohoto textu je čtenáře stručně seznámit s nástroji algebraické teorie formálních jazyků a s některými významnými příklady jejich aplikací. Přestože pro jeho čtení je postačující znalostí ovládnutí základních matematických pojmů a konstrukcí (pologrupy, grupy, variety algeber, rovnostní logika, metrické prostory), ke správnému pochopení jeho významu je vhodné se dříve seznámit se základy teorie formálních jazyků.

V tomto textu se budeme věnovat především jazykům regulárním, tvořícím nejnižší úroveň Chomského hierarchie. Jedná se o jazyky, pro něž existuje konečný mechanismus, který umí o každém slově rozhodnout, zda do jazyka patří nebo ne. Hlavním cílem většiny výsledků v teorii regulárních jazyků je algoritmické rozhodování vlastností regulárních jazyků a hledání efektivních charakterizací významných tříd těchto jazyků. Například nás bude zajímat, jak zjistit, jestli je možné daný jazyk získat pomocí nějakým způsobem omezených racionálních výrazů, třeba s omezením na používané operace nebo s omezením na hloubku zanoření operací ve výrazu.

Regulární jazyky bývají nejčastěji popisovány pomocí konečných automatů nebo regulárních výrazů, ovšem jsou známy i další ekvivalentní přístupy, z nichž každý má své výhody a nevýhody. Jednou z možností je použití formulí monadické logiky druhého řádu; tento fakt patří k hlavním motivacím pro rozvoj teorie regulárních jazyků jako cesty k nalezení nástrojů použitelných k algoritmické manipulaci s logickými formulemi. Jinou možností je použití homomorfismů do konečných pologrup, což umožňuje využít ke studiu jazyků jak strukturní teorii konečných pologrup, tak techniky univerzální algebry, především identity.

Začátky studia pologrup spadají již do 20. let 20. století. Formální jazyky se staly středem zájmu ve vznikající teoretické informatice v 50. letech. Prvním výsledkem, který ukázal význam teorie pologrup pro studium formálních jazyků byla Schützenbergerho charakterizace star-free jazyků z roku 1965. Toto zjištění bylo brzy potvrzeno množstvím dalších výsledků představujících úzké vazby mezi konečnými pologrupami a regulárními jazyky. Obecný rámec pro tyto výsledky vytvořil v roce 1976 Samuel Eilenberg zavedením pojmu variety jazyků. Od té doby tvoří regulární jazyky a konečné pologrupy v podstatě jednu společnou teorii, přičemž mnoho výsledků o jazycích umíme dokázat pouze pomocí pologrup a naopak, při studiu vlastností konečných pologrup se dobře uplatňují kombinatorické argumenty o jim odpovídajících jazycích. Přestože vzniklá

teorie je dnes již poměrně rozsáhlá a obsahuje mnoho hlubokých výsledků, některé problémy, které motivovaly její vznik, zůstávají stále nevyřešené.

První kapitola se zabývá otázkou, jaké podmnožiny pologrup lze popsat pomocí obvyklých mechanismů, které v případě monoidů slov popisují regulární jazyky. Každý z uvažovaných mechanismů se ukáže být alternativním popisem buď deterministického, nebo nedeterministického, konečného automatu. Vztah mezi deterministickou a nedeterministickou verzí různých výpočetních modelů popisujících množiny slov patří k základním otázkám teoretické informatiky. Zatímco pro některé modely je snadné zjistit, že deterministická a nedeterministická verze mají stejnou sílu (třeba pro konečné automaty nebo Turingovy stroje) a pro jiné je dobře známo, že nedeterministické stroje jsou silnější (třeba v případě zásobníkových automatů), například pro složitostní třídy určené asymptotickou časovou náročností výpočtu se často jedná o obtížné otevřené problémy (třeba slavný $P \stackrel{?}{=} NP$ problém). Uvažujeme-li místo konečných slov nad konečnou abecedou prvky libovolné konečně generované pologrupy, ukazuje se, že nedeterministické konečné automaty mají obvykle větší vyjadřovací sílu než automaty deterministické.

Druhá kapitola se věnuje klasické strukturní teorii pologrup, se zaměřením na výsledky použitelné v teorii formálních jazyků. Pro každý z prezentovaných výsledků je předvedeno jeho použití k důkazu nějakého významného výsledku o regulárních jazycích.

Ve třetí kapitole je vybudována obecná teorie pseudovariet konečných pologrup, tedy tříd uzavřených na součiny, podpologrupy a homomorfní obrazy, včetně jejich rovnostního popisu pomocí pseudoidentit; není možné použít obvyklou Birkhoffovu teorii variet, neboť ta funguje pouze tehdy, máme-li k dispozici nekonečné součiny algeber. Poté je dokázána jednoznačná korespondence variet pologrup s varietami regulárních jazyků a je uveden stručný přehled zajímavých příkladů této korespondence, které se podařilo dokázat.

V poslední kapitole je rozpoznávání jazyků homomorfismy do konečných pologrup zobecněno na třídu dobře částečně uspořádaných pologrup, o které je známo, že rovněž rozpoznává pouze regulární jazyky. Je ukázáno, jak lze tohoto faktu využít k získání nekonstruktivních důkazů regularity jazyků.

Většina látky popisované v tomto textu je obsažena v dosud nedokončené knize Jean-Érica Pina [6], jejíž stará verze vyšla v roce 1986 [5]. Části textu označené hvězdičkou představují stručnou informaci o zajímavých výsledcích úzce souvisejících s probíranou látkou a je možné je při čtení zcela vynechat. Argumenty, o kterých se autor domnívá, že je pro čtenáře užitečné, aby si je doplnil sám, jsou uvedena jako cvičení. Vyřešení většiny těchto cvičení by mělo být pro pečlivého čtenáře přímočarou záležitostí; hvězdičkou jsou (***budou) označena cvičení, k jejichž vyřešení je třeba nějaká nová myšlenka.

Kapitola 1

Rozpoznatelné a racionální podmnožiny pologrup

V první kapitole zavedeme základní pojem regulárního jazyka. Přitom budeme postupovat obecněji, než je obvyklé v úvodním kurzu formálních jazyků. Budeme totiž považovat za jazyky nejen množiny slov, ale množiny prvků libovolné pevně zvolené pologrupy. Zjistíme, že při tomto zobecnění ztratíme ekvivalenci standardních definic regulárních jazyků pomocí deterministických a nedeterministických automatů. Pro vzniklé třídy podmnožin proto budeme používat jiné termíny: rozpoznatelné a racionální podmnožiny. O regulárních množinách tedy budeme mluvit pouze v případě, že se tyto dvě třídy shodují.

Většinu materiálu prezentovaného v této kapitole je možné v nějaké podobě najít detailněji zpracované buď v bakalářské práci Marka Filakovského [3] nebo v knize Jacquesa Sakarovitche [8]. S výjimkou části 1.6 tuto kapitolu pokrývá také kniha Jean-Érica Pina [6].

1.1 Pologrupy a monoidy

Připomeňme, že *pologrupa* je množina S opatřená asociativní binární operací $\cdot : S \times S \to S$. Přitom, pokud je z kontextu jasné, se kterou operací pracujeme, mluvíme často jednoduše o pologrupě S místo $(S, \cdot)$ a píšeme součin prvků $x, y \in S$ místo $x \cdot y$ zkráceně xy. *Monoidem* rozumíme pologrupu obsahující neutrální prvek, který obvykle značíme 1.

Každou pologrupu S lze snadno vnořit do monoidu přidáním nového neutrálního prvku. Přitom existuje až na izomorfismus nejmenší monoid obsahující S, který se obvykle značí S^1. Pokud S není sama monoidem, potom $S^1 = S \cup \{1\}$, kde $1 \notin S$ je nový prvek a násobení tímto prvkem je definováno předpisem $x \cdot 1 = 1 \cdot x = x$ pro všechna $x \in S^1$.

Říkáme, že prvek e pologrupy S je *idempotentní*, jestliže platí $e \cdot e = e$. Množina všech

idempotentních prvků pologrupy S se obvykle značí $E(S)$. Prvek $0 \in S$ splňující pro všechna $x \in S$ rovnosti $0 \cdot x = x \cdot 0 = 0$ se nazývá *nulový prvek* pologrupy S. Připomeňme, že jak neutrální, tak nulový, prvek pologrupy jsou určeny jednoznačně, avšak dalších idempotentních prvků může pologrupa obsahovat libovolně mnoho.

Homomorfismus pologrupy S do pologrupy T je zobrazení $\varphi\colon S \to T$ splňující $\varphi(xy) = \varphi(x)\varphi(y)$ pro všechny prvky x a y z S. V případě, že S i T jsou monoidy a homomorfismus $\varphi\colon S \to T$ navíc zobrazuje neutrální prvek S na neutrální prvek T, mluvíme o *homomorfismu monoidů*. *Kongruencí* pologrupy S myslíme relaci ekvivalence $\rho \subseteq S \times S$ takovou, že pro všechna $x, x', y, y' \in S$ splňující $x \,\rho\, x'$ a $y \,\rho\, y'$ platí rovněž $xy \,\rho\, x'y'$. Tato podmínka nám umožňuje korektně definovat na množině S/ρ strukturu pologrupy násobením pomocí reprezentantů, přičemž přirozená projekce $\nu\colon S \to S/\rho$ je surjektivním homomorfismem. Opačně, jádro libovolného homomorfismu $\ker(\varphi) = \{(x,y) \in S \times S \mid \varphi(x) = \varphi(y)\}$ je kongruencí pologrupy S a obraz S v homomorfismu φ je přirozeně izomorfní pologrupě $S/\ker(\varphi)$. *Antihomomorfismem* $\varphi\colon S \to T$ rozumíme zobrazení splňující $\varphi(xy) = \varphi(y)\varphi(x)$ pro všechna $x, y \in S$.

Říkáme, že pologrupa T *dělí* pologrupu S, jestliže T je homomorfním obrazem nějaké podpologrupy S.

1.1.1 Volné monoidy

Abecedou myslíme nějakou konečnou množinu A, jejímž prvkům říkáme *písmena*. *Slovo* nad abecedou A je konečná posloupnost písmen z A. V následujícím textu se budeme nejčastěji potkávat s monoidem A^* tvořeným všemi slovy nad danou abecedou A; operací v tomto monoidu je zřetězení slov, tedy pro $u, v \in A^*$ máme $u \cdot v = uv$. Pologrupa, která vznikne z tohoto monoidu odebráním prázdného slova ε, se značí A^+. Uvědomte si, že pokud abeceda A obsahuje jediné písmeno, je monoid A^* izomorfní monoidu nezáporných celých čísel s operací sčítání $(\mathbb{N}_0, +)$. Pro libovolné přirozené číslo n značíme $A^{\geq n}$ množinu všech slov nad A délky alespoň n, tedy $A^{\geq n} = A^n \cdot A^*$.

V teorii formálních jazyků se většinou studují podmnožiny monoidu A^*, nazývané obvykle *jazyky*. My se v této kapitole neomezíme jen na tento monoid, ale definujeme rozpoznatelné podmnožiny obecně v libovolné pologrupě.

Připomeňme si nejdůležitější vlastnost monoidu všech slov: zvolíme-li libovolně obrazy písmen v jakémkoli monoidu M, toto přiřazení lze právě jedním způsobem rozšířit na homomorfismus monoidu A^* do monoidu M. Totéž tvrzení platí pro homomorfismy pologrupy A^+ do libovolné pologrupy S.

Při práci se slovy budeme používat následující pojmy. Množinu všech písmen abecedy A vyskytujících se ve slově $w \in A^*$ budeme značit $\mathrm{alph}(w)$. Délku slova w budeme značit $|w|$ a pro počet výskytů písmene $a \in A$ ve slově w budeme používat označení $|w|_a$. Slovo $u \in A^*$ nazýváme *faktorem* (*prefixem*, *sufixem*) slova $v \in A^*$, jestliže existují slova $w, x \in A^*$ taková, že $v = wux$ ($v = uw$, $v = wu$). Slovo $a_1 \ldots a_n$, kde $a_1, \ldots, a_n \in A$, nazýváme (roztroušeným) *podslovem* slova $v \in A^*$, jestliže existují slova $u_0, \ldots, u_n \in A^*$

taková, že $v = u_0 a_1 u_1 \dots a_n u_n$.

1.1.2 Pologrupy transformací

Pologrupy transformací dané množiny hrají v teorii pologrup stejnou roli jako grupy permutací v teorii grup. Z pohledu teorie jazyků spočívá jejich význam v tom, že zprostředkovávají vazbu mezi deterministickými automaty a pologrupami, jelikož přechodová funkce deterministického automatu vlastně udává pro každé slovo nějakou transformaci množiny všech stavů tohoto automatu.

Je-li Q libovolná množina, nazýváme množinu $\mathscr{T}(Q)$ všech zobrazení Q do Q spolu s operací skládání *úplným transformačním monoidem*. Přitom budeme zobrazení aplikovat zleva, tedy $(f \circ g)(q) = f(g(q))$.

Tvrzení 1.1. *Nechť S je libovolná pologrupa. Potom zobrazení $\varphi\colon S \to \mathscr{T}(S^1)$ dané předpisem $\varphi(x)(y) = x \cdot y$, pro všechna $x \in S$ a $y \in S^1$, je vložením pologrupy S do monoidu $\mathscr{T}(S^1)$. Mimo jiné je tedy každá (konečná) pologrupa izomorfní nějaké pologrupě transformací (konečné množiny).*

Cvičení 1.2.

1. Dokažte předchozí tvrzení.
2. Dejte příklad pologrupy S, která není izomorfní žádné podpologrupě pologrupy $\mathscr{T}(S)$.

Často používaným zobecněním transformací jsou parciální transformace, které odpovídají neúplným deterministickým automatům. Přitom pologrupa všech parciálních transformací $\mathscr{P}\mathscr{T}(Q)$ je izomorfní podpologrupě pologrupy $\mathscr{T}(Q \cup \{s\})$, kde s je nový prvek odpovídající stavu automatu, ze kterého nevede žádný přechod jinam (sink state). Příslušné zobrazení $\varphi\colon \mathscr{P}\mathscr{T}(Q) \to \mathscr{T}(Q \cup \{s\})$ posílá každou parciální transformaci f na transformaci, která vznikne dodefinováním $\varphi(f)(q) = s$ pro všechna $q \in Q$, na nichž není f definováno.

Dalším zobecněním parciálních transformací se přirozeně dostaneme k pologrupám binárních relací. V důkazu věty 1.42 uvidíme, že v případě rozpoznávání standardních jazyků (ve volném monoidu) odpovídají pologrupy relací nedeterministickým automatům. Ovšem v případě jiných než volných monoidů kompozice binárních relací obecně neodpovídá sémantice nedeterministických automatů, a proto pologrupy binárních relací vazbu mezi nedeterministickými automaty a pologrupami zprostředkovat nemohou.

1.2 Rozpoznatelné podmnožiny

1.2.1 Rozpoznávání homomorfismy

Nechť S je nyní libovolná pevně zvolená pologrupa. Říkáme, že podmnožina $L \subseteq S$ je *rozpoznávána* homomorfismem $\varphi\colon S \to T$ do nějaké pologrupy T, jestliže existuje

podmnožina $F \subseteq T$ taková, že $L = \varphi^{-1}(F)$. V této situaci rovněž někdy říkáme, že L je rozpoznávána pologrupou T. Být rozpoznávána jistou pologrupou tedy znamená být vzorem nějaké podmnožiny této pologrupy v nějakém homomorfismu. Podstata této definice je v tom, že za S můžeme brát nějakou velkou nekonečnou pologrupu a za T malou (například konečnou) pologrupu. Potom totiž můžeme na homomorfismus φ nahlížet jako na malý nástroj sloužící k popisu velkých množin. Nejčastější situací, kdy se tato definice používá, je případ, kdy pologrupa S je konečně generovaná a pologrupa T je konečná. Potom nám totiž k zadání libovolného takového homomorfismu φ stačí uvést konečnou informaci: popsat pologrupu T a určit obrazy všech generátorů pologrupy S.

Všimněte si, že za množinu F v definici rozpoznávání můžeme vždy brát množinu $\varphi(L)$, takže homomorfismus φ rozpoznává L právě tehdy, když $\varphi^{-1}(\varphi(L)) = L$. Rozpoznávání množiny L homomorfismem φ tedy znamená, že pro zjištění, zda prvek $x \in S$ patří do L, stačí tento prvek zobrazit pomocí φ do T a podívat se, jestli $\varphi(x)$ patří do $\varphi(L)$.

Říkáme, že podmnožina L pologrupy S je *rozpoznatelná* (recognisable) podmnožina S, jestliže existuje homomorfismus $\varphi\colon S \to T$ do nějaké konečné pologrupy T, který rozpoznává L. Množina všech rozpoznatelných podmnožin pologrupy S se obvykle značí $\mathrm{Rec}(S)$.

Předchozí definice lze analogicky formulovat i pro monoidy a v následujícím textu budeme vždy používat tu definici, která je v dané situaci vhodnější. Většinou si tedy vystačíme s rozpoznáváním monoidy, pouze pokud se budeme zabývat některými třídami jazyků neobsahujících prázdné slovo, budeme muset použít obecnější definici pro pologrupy.

Příklad 1.3. Ukažme si, že libovolná konečná podmnožina $L \subseteq A^*$ je rozpoznatelná. Za tímto účelem uvažme množinu $\mathrm{fac}(L)$ všech faktorů slov patřících do L a přidejme k ní nový prvek 0. Na vzniklé množině $M = \mathrm{fac}(L) \cup \{0\}$ definujeme násobení předpisem

$$u \cdot v = \begin{cases} uv, & \text{jestliže } uv \in \mathrm{fac}(L), \\ 0, & \text{jinak,} \end{cases}$$

a $u \cdot 0 = 0 \cdot v = 0 \cdot 0 = 0$ pro libovolná slova $u, v \in \mathrm{fac}(L)$. Nyní se již snadno ověří, že takto definované násobení korektně zadává na M strukturu monoidu a že přirozený homomorfismus, který zobrazuje každé slovo z $\mathrm{fac}(L)$ na sebe a ostatní slova monoidu A^* na nulu, rozpoznává jazyk L.

Cvičení 1.4.

1. Ukažte, že v grupě $(\mathbb{Z}, +)$ jsou rozpoznatelné právě podmnožiny, které jsou sjednocením tříd modulo n pro nějaké přirozené číslo n. Uvědomte si, že tedy jedinou rozpoznatelnou konečnou podmnožinou $\mathbb{Z}$ je prázdná množina, která je rozpoznatelnou podmnožinou každé pologrupy.

2. Ukažte, že v pologrupě $(\mathbb{N},+)$ jsou rozpoznatelné právě ty množiny L, pro které existují přirozená čísla m a n taková, že příslušnost čísel větších než m do L závisí pouze na jejich zbytku modulo n. Těmto množinám přirozených čísel se říká *ultimativně periodické*.

Říkáme, že relace ekvivalence ρ na množině S *nasycuje* podmnožinu $L \subseteq S$, jestliže je L sjednocením tříd rozkladu S/ρ, tedy jestliže pro všechna $x, y \in S$ splňující $x \,\rho\, y$ a $x \in L$ platí rovněž $y \in L$. Díky korespondenci mezi homomorfismy a kongruencemi snadno vidíme, že $L \subseteq S$ je rozpoznatelná právě tehdy, když existuje kongruence pologrupy S, která má konečný index a nasycuje L.

Obvyklým nástrojem, který v jednoduchých případech umožňuje dokázat, že daná podmnožina není rozpoznatelná, je takzvané lemma o vkládání (pumping lemma):

Lemma 1.5 (Michael O. Rabin a Dana S. Scott, 1959)**.** *Nechť L je rozpoznatelná podmnožina pologrupy S. Potom existuje přirozené číslo m takové, že pro libovolné $n \geq m$ a pro libovolné prvky $x_1, \ldots, x_n \in S$ splňující $x_1 \cdots x_n \in L$ existují $i, j \in \mathbb{N}$, $1 \leq i \leq j \leq m$, takové, že prvek $x_1 \cdots x_{i-1}(x_i \cdots x_j)^k x_{j+1} \cdots x_n$ náleží do L pro všechna $k \in \mathbb{N}_0$.*

Cvičení 1.6. Dokažte předchozí lemma. Použijte toto lemma k důkazu, že jazyk $\{a^n b^n \mid n \in \mathbb{N}\}$ nad abecedou $A = \{a, b\}$ není rozpoznatelný.

1.2.2 Deterministické automaty

Definice rozpoznávání pomocí pologrup je výhodná, pokud chceme při práci s jazyky využít abstraktní algebraické postupy. Pokud ovšem chceme názorně popsat nějaký rozpoznatelný jazyk, použijeme většinou deterministické automaty. Podstatnou výhodou automatu je nejen to, že na rozdíl od binární operace se dá znázornit jako orientovaný multigraf, ale rovněž jeho velikost, neboť nejmenší rozpoznávající pologrupa je v některých případech i více než exponenciálně větší než automat popisující stejný jazyk.

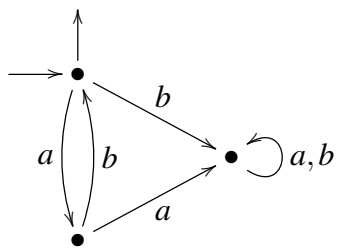
Definice 1.7. *Akcí* pologrupy S na množině Q rozumíme zobrazení $\delta\colon Q \times S \to Q$ splňující pro všechna $x, y \in S$ a $q \in Q$ podmínku $\delta(\delta(q,x),y) = \delta(q,xy)$. Chápeme-li pologrupu S jako monoid, vyžadujeme navíc, aby platilo $\delta(q,1) = q$ pro všechna $q \in Q$.

Na akci pologrupy S na množině Q se můžeme ekvivalentně dívat jako na antihomomorfismus φ pologrupy S do pologrupy $\mathscr{T}(Q)$, přičemž $\varphi(x)(q) = \delta(q,x)$. Přitom transformaci $\varphi(x)$ obvykle značíme δ_x, tedy platí $\delta_{xy} = \delta_y \circ \delta_x$. Dodatečný požadavek v případě monoidu potom odpovídá tomu, že se má jednat o antihomomorfismus monoidů, tedy že se neutrální prvek S zobrazí na identickou transformaci. Všimněte si, že akce neodpovídají homomorfismům, ale antihomomorfismům, protože zobrazení aplikujeme na argumenty zleva, tedy v pořadí odprava doleva, kdežto slova čteme zleva doprava. Proto také bývá zvykem při studiu akcí předpokládat, že zobrazení se aplikují zprava, přičemž, je-li jasné, o kterou akci se jedná, píše se stručně $q \cdot x$ místo $\delta_x(q)$.

Definice 1.8. Nechť S je pologrupa. *Deterministickým automatem* nad S rozumíme uspořádanou pětici $\mathscr{A} = (S, Q, \delta, q_0, F)$, kde Q je nějaká množina *stavů*, $\delta\colon Q \times S \to Q$ je akce S na Q, $q_0 \in Q$ je *počáteční stav* a $F \subseteq Q$ je množina *koncových stavů*. Říkáme, že automat $\mathscr{A}$ je *konečný*, jestliže množina Q je konečná.

Podmnožinu $L(\mathscr{A}) = \{x \in S \mid \delta(q_0, x) \in F\}$ pologrupy S nazýváme podmnožinou *přijímanou* automatem $\mathscr{A}$.

Všimněte si, že akci pologrupy na množině stačí zadat pouze pro generátory pologrupy S. Proto v případě konečně generované pologrupy S stačí k popsání libovolného konečného automatu nad S zadat pouze konečnou informaci v podobě orientovaného multigrafu jako na obrázku 1.1. Je ovšem třeba mít na paměti, že na rozdíl od obvyklého případu automatů nad pologrupou A^+ obecně ne každý takový graf určuje deterministický automat. V případě pologrupy A^+ je tomu tak proto, že se jedná o volnou pologrupu nad abecedou A, a proto každé zobrazení prvků A do $\mathscr{T}(Q)$ určuje homomorfismus pologrup.



Obrázek 1.1: Orientovaný multigraf zadávající deterministický automat nad A^+ přijímající jazyk $(ab)^+$, ale nezadávající deterministický automat nad monoidem $\mathbb{N}_0 \times \mathbb{N}_0$ s generátory $a = (1,0)$ a $b = (0,1)$, neboť $\delta(q_0, ab) \neq \delta(q_0, ba)$.

Tak jako rozpoznávání podmnožin pologrupy S pomocí homomorfismů odpovídá nasycování kongruencemi pologrupy S, přijímání automaty odpovídá nasycování jednostrannými kongruencemi pologrupy S. Říkáme, že relace ekvivalence ρ na pologrupě S je *pravou* (respektive *levou*) *kongruencí* S, jestliže pro všechna $x, y, z \in S$ splňující $x \,\rho\, y$ platí rovněž $xz \,\rho\, yz$ (respektive $zx \,\rho\, zy$).

Cvičení 1.9. Dokažte, že relace ekvivalence ρ na S je kongruencí S právě tehdy, když je současně pravou i levou kongruencí.

Je-li ρ pravá kongruence S nasycující L, potom můžeme na množině stavů $S/\rho \cup \{q_0\}$, kde $q_0 \notin S/\rho$ je nový prvek sloužící jako počáteční stav, definovat strukturu deterministického automatu přijímajícího L předpisy

$$\delta(q_0, y) = [y]_\rho \qquad \text{a} \qquad \delta([x]_\rho, y) = [xy]_\rho,$$

pro všechna $x, y \in S$, a položením $F = L/\rho$. Všimněte si, že nový počáteční stav potřebujeme k množině S/ρ přidat pouze v případě, že S/ρ není monoid; pokud je totiž S/ρ monoid, můžeme za počáteční stav vzít jeho neutrální prvek.

Naopak, každý deterministický automat přijímající L definuje na S pravou kongruenci nasycující L předpisem

$$x \, \rho \, y \quad \Longleftrightarrow \quad \delta(q_0, x) = \delta(q_0, y).$$

Snadno se tedy vidí, že podmnožina je přijímaná konečným deterministickým automatem právě tehdy, když je nasycována pravou kongruencí konečného indexu. Podobně se nahlédne, že podmnožina pologrupy S je nasycována levou kongruencí konečného indexu právě tehdy, když je přijímaná deterministickým automatem nad pologrupou antiizomorfní k S (násobení v této pologrupě je dáno předpisem $x \bullet y = y \cdot x$; v případě volného monoidu A^* se tedy vlastně jedná o automat přijímající převrácení daného jazyka nad A^*).

Jak si ovšem nyní ukážeme, tyto dvě situace jsou ve skutečnosti obě ekvivalentní s rozpoznatelností dané podmnožiny. Nejprve si všimněme, že rozpoznatelnost L znamená existenci kongruence konečného indexu nasycující L a tedy podle cvičení 1.9 i existenci pravé a levé kongruence.

Na druhou stranu, je-li množina rozpoznávaná konečným deterministickým automatem, můžeme zkonstruovat homomorfismus do konečné pologrupy, který tuto množinu rozpoznává. Touto pologrupou je takzvaná přechodová pologrupa automatu. Chápeme-li akci δ jako antihomomorfismus S do $\mathscr{T}(Q)$, je *přechodovou pologrupou* (transition semigroup) automatu $\mathscr{A}$ právě obraz pologrupy S v tomto antihomomorfismu, tedy množina zobrazení $\{\delta_x \mid x \in S\}$ spolu s kompozicí zobrazení. Potom antihomomorfismus přiřazující prvku x transformaci δ_x rozpoznává L množinou $\{\delta_x \mid \delta_x(q_0) \in F\}$.

1.2.3 Rozpoznávání uspořádanými pologrupami

V tomto odstavci zobecníme pojem rozpoznávání pologrupami na rozpoznávání pologrupami vybavenými kompatibilním uspořádáním. Cílem tohoto postupu, poprvé použitého až Jean-Éricem Pinem v roce 1995, je především získání nástroje pro jemnější klasifikaci rozpoznatelných jazyků, než umožňují pologrupy bez uspořádání.

Připomeňme, že předuspořádáním na množině T myslíme reflexivní a tranzitivní binární relaci na T. Říkáme, že předuspořádání $\leq$ na pologrupě T je *monotónní*, jestliže pro libovolná $x, x', y, y' \in T$ z platnosti $x \leq x'$ a $y \leq y'$ plyne $x \cdot y \leq x' \cdot y'$. Uvědomte si, že kongruenci můžeme definovat jako monotónní relaci ekvivalence.

Pologrupa vybavená monotónním uspořádáním se nazývá *uspořádaná pologrupa*. Povšimněme si, že každou pologrupu lze chápat jako uspořádanou pologrupu, pokud za příslušné uspořádání vezmeme relaci identity $=$. Pokud bude příslušné monotónní uspořádání $\leq$ jasné z kontextu, budeme mluvit stručně o uspořádané pologrupě T místo $(T, \leq)$.

Cvičení 1.10. Dokažte, že na každé konečné grupě je identita jediným monotónním uspořádáním.

Cvičení 1.11. Ukažte, že na tříprvkovém polosvazu, který je řetězcem, (jinými slovy, tříprvkovém idempotentním monoidu s nulovým prvkem) existuje právě 13 monotónních uspořádání.

Říkáme, že podmnožina L pologrupy S je *rozpoznávána* homomorfismem $\varphi\colon S \to T$ do nějaké uspořádané pologrupy T, jestliže existuje podmnožina $F \subseteq T$ nahoru uzavřená vzhledem k uspořádání $\leq$ a taková, že $L = \varphi^{-1}(F)$. Všimněte si, že v případě, kdy pologrupa T je uspořádaná relací identity, dostáváme přesně původní pojem rozpoznávání.

Tak jako homomorfismy do obyčejných pologrup odpovídají kongruencím pologrupy S, homomorfismy do uspořádaných pologrup odpovídají monotónním předuspořádáním na S. Přesněji, každý homomorfismus $\varphi\colon S \to T$ do uspořádané pologrupy T indukuje monotónní předuspořádání na S předpisem

$$x \leq_\varphi y \quad \Longleftrightarrow \quad \varphi(x) \leq \varphi(y).$$

Podmnožina $L \subseteq S$ je potom rozpoznávaná homomorfismem φ právě tehdy, když je nahoru uzavřená vzhledem k $\leq_\varphi$.

Opačně, pro libovolné monotónní předuspořádání $\preceq$ na S můžeme uvážit jemu příslušející relaci ekvivalence $\sim$ definovanou předpisem

$$x \sim y \quad \Longleftrightarrow \quad x \preceq y \text{ a } y \preceq x.$$

Tato relace ekvivalence je kongruencí pologrupy S a předpis

$$[x]_\sim \leq [y]_\sim \quad \Longleftrightarrow \quad x \preceq y$$

definuje monotónní uspořádání na pologrupě $S/{\sim}$. Příslušným homomorfismem je potom přirozená projekce na kvocient $\nu\colon S \to S/{\sim}$.

Z právě popsaných konstrukcí vidíme, že homomorfismy do konečných uspořádaných pologrup odpovídají právě monotónním předuspořádáním, jimž příslušející relace ekvivalence má konečný index.

Jediným rozdílem oproti rozpoznávání neuspořádanou pologrupou tedy je, že uspořádání nám vybírá jen některé z jazyků, které jsou rozpoznávány pologrupou bez uspořádání. Tímto nám umožňuje jemnější klasifikaci rozpoznatelných jazyků než prosté rozpoznávání homomorfismy do konečných pologrup.

Poznámka 1.12. V původní Pinově definici rozpoznávání uspořádanými pologrupami se pracuje s dolů uzavřenými podmnožinami F, namísto nahoru uzavřených. V tomto textu budeme používat nahoru uzavřené podmnožiny kvůli konzistenci s tradičním pojmem dobrých předuspořádání, kterému se budeme věnovat v kapitole 4. Při studiu literatury používající opačnou notaci je tedy třeba mít na paměti, že všechny nerovnosti je třeba otáčet.

1.2.4 Minimální automat a syntaktický homomorfismus

Nyní si ukážeme, že jak mezi automaty, tak mezi homomorfismy, které popisují daný jazyk, existuje vždy jistý kanonický. Jejich definice je založena na pojmu kontextu prvku v podmnožině pologrupy. Je-li S pologrupa, $x \in S$ nějaký její prvek a $L \subseteq S$ její podmnožina, potom množinu všech (oboustranných) *kontextů* prvku x v L definujeme jako

$$\mathrm{C}_L(x) = \{(s,t) \mid s,t \in S^1,\ sxt \in L\}.$$

Podobně zavedeme pojmy levého a pravého kontextu:

$$\mathrm{C}_L^\ell(x) = \{s \in S^1 \mid sx \in L\},$$
$$\mathrm{C}_L^r(x) = \{t \in S^1 \mid xt \in L\}.$$

Pro pevně zvolenou podmnožinu $L \subseteq S$ obdržíme porovnáváním kontextů různých prvků pologrupy S monotónní předuspořádání na S, které se nazývá *syntaktické monotónní předuspořádání* podmnožiny L. Definujeme tedy

$$x \leq_L y \quad \Longleftrightarrow \quad \mathrm{C}_L(x) \subseteq \mathrm{C}_L(y)$$

pro libovolné prvky $x, y \in S$.

Cvičení 1.13. Ověřte, že $\leq_L$ je skutečně monotónní předuspořádání S.

Jak ukazuje následující lemma, takto definované předuspořádání je největší monotónní předuspořádání S, vzhledem ke kterému je množina L nahoru uzavřená.

Lemma 1.14. *Nechť $L \subseteq S$ je libovolná podmnožina a $\leq$ libovolné monotónní předuspořádání na S. Potom L je nahoru uzavřená vzhledem k $\leq$ právě tehdy, když $\leq \subseteq \leq_L$.*

Cvičení 1.15. Dokažte předchozí lemma. Všimněte si, že toto tvrzení by neplatilo, pokud bychom v definici $\mathrm{C}_L(x)$ měli místo S^1 přímo pologrupu S.

Opět můžeme podobně zavést příslušné jednostranné analogie, *levé a pravé syntaktické předuspořádání*:

$$x \leq_L^\ell y \iff \mathrm{C}_L^\ell(x) \subseteq \mathrm{C}_L^\ell(y),$$
$$x \leq_L^r y \iff \mathrm{C}_L^r(x) \subseteq \mathrm{C}_L^r(y)$$

Kongruence $\sim_L$ pologrupy S příslušející syntaktickému předuspořádání se nazývá *syntaktická kongruence* podmnožiny $L \subseteq S$. Tedy platí $x \sim_L y$ právě tehdy, když $\mathrm{C}_L(x) = \mathrm{C}_L(y)$. Kvocient pologrupy S podle této kongruence $S_L = S/{\sim_L}$ se nazývá *syntaktická pologrupa* podmnožiny L. Přirozené projekci $\varphi_L \colon S \to S_L$ se potom říká *syntaktický homomorfismus*. Protože předuspořádání $\leq_L$ přirozeně indukuje monotónní uspořádání

na pologrupě S_L, můžeme mluvit rovněž o *syntaktické uspořádané pologrupě* podmnožiny L. Pracujeme-li místo s pologrupami s monoidy, mluvíme o *syntaktickém monoidu* M_L.

Syntaktická kongruence má podobnou univerzální vlastnost jako syntaktické předuspořádání, tedy je největší mezi všemi kongruencemi S nasycujícími L.

Lemma 1.16. *Nechť $L \subseteq S$ je libovolná podmnožina a $\sim$ libovolná kongruence pologrupy S. Potom $\sim$ nasycuje L právě tehdy, když $\sim \subseteq \sim_L$.*

Cvičení 1.17. Dokažte předchozí lemma.

Z korespondence mezi homomorfismy a kongruencemi a z předchozího lemmatu vyplývá, že je-li T libovolná pologrupa rozpoznávající podmnožinu $L \subseteq S$, potom syntaktická pologrupa podmnožiny L je homomorfním obrazem podpologrupy pologrupy T. Proto je S_L mezi pologrupami rozpoznávajícími L nejmenší vzhledem k dělitelnosti. Tuto vlastnost je možné stručně vyjádřit komutativním diagramem

$$\begin{tikzcd} S \arrow[r, two heads, "\varphi"] \arrow[dr, two heads, "\varphi_L"'] & T \arrow[d, "\exists!\psi"] \\ & S_L \end{tikzcd} \tag{1.1}$$

kde φ je libovolný surjektivní homomorfismus rozpoznávající L.

Chceme-li podobným způsobem vytvořit kanonický automat podmnožiny L pologrupy S, uvážíme pravé syntaktické předuspořádání L na S a jemu příslušející pravou kongruenci $\sim_L^r$ pologrupy S. Nyní vytvoříme na množině $S/\sim_L^r$ deterministický automat postupem popsaným na straně 8, přičemž nový počáteční stav přidáme pouze tehdy, když neexistuje prvek $x \in S$ splňující $\mathrm{C}_L^r(x) = L$; pokud takový prvek existuje, zvolíme za počáteční stav $[x]_{\sim_L^r}$. Vzniklý automat přijímá L a nazývá se *minimální automat* podmnožiny L pologrupy S. Analogicky předchozím úvahám můžeme ověřit, že minimální automat je nejmenší mezi deterministickými automaty přijímajícími L, tedy že je obrazem nějakého podautomatu každého automatu přijímajícího L.

Cvičení 1.18. Dokažte, že přechodová pologrupa minimálního automatu podmnožiny $L \subseteq S$ je přirozeně antiizomorfní syntaktické pologrupě podmnožiny L.

Cvičení 1.19. Dokažte, že jazyk $L \subseteq A^+$ je rozpoznávaný konečnou grupou právě tehdy, když jeho minimální automat je konečný a duálně deterministický, tedy jestliže v něm neexistují dva různé stavy p a q takové, že pro nějaké $a \in A$ platí $\delta(p,a) = \delta(q,a)$.

Ukažme si nyní dva jednoduché příklady, jak se vlastnosti jazyka nebo vztah daného slova k tomuto jazyku promítají do syntaktického monoidu.

Příklad 1.20.

1. Máme-li jazyk L nad určitou abecedou a přidáme k této abecedě několik nových písmen, která se ve slovech jazyka L nevyskytují, budou tato nová písmena vždy reprezentovat nulový prvek syntaktické pologrupy L nad rozšířenou abecedou. Přitom nová písmena reprezentují nulový prvek původní syntaktické pologrupy právě tehdy, když některé slovo nad původní abecedou není faktorem žádného slova jazyka L; v opačném případě nová syntaktická pologrupa vznikne z původní syntaktické pologrupy přidáním nového nulového prvku. Z tohoto příkladu ihned vidíme, že některé algebraické vlastnosti syntaktického monoidu závisejí na tom, nad kterou abecedou daný jazyk uvažujeme.

2. Pro $x \in S$ je prvek $\varphi_L(x)$ pologrupy S_L idempotentní právě tehdy, když libovolné zvyšování počtu sousedících kopií x v součinu nějakých prvků pologrupy S neovlivňuje příslušnost tohoto součinu do L, tedy pro všechna $y, z \in S^1$ a $n \in \mathbb{N}$ platí $yxz \in L$ právě tehdy, když $yx^n z \in L$.

Následující věta shrnuje dosud dokázané poznatky o možných charakterizacích rozpoznatelných podmnožin dané pologrupy.

Věta 1.21. *Nechť S je pologrupa a $L \subseteq S$ její libovolná podmnožina. Potom následující podmínky jsou ekvivalentní.*

1. *podmnožina L je rozpoznávaná homomorfismem do konečné pologrupy,*
2. *podmnožina L je rozpoznávaná homomorfismem do konečné uspořádané pologrupy,*
3. *podmnožina L je přijímaná konečným deterministickým automatem,*
4. *podmnožina L je nahoru uzavřená vzhledem k nějakému monotónnímu předuspořádání pologrupy S, které má jen konečně mnoho tříd ekvivalentních prvků,*
5. *pologrupa S_L je konečná,*
6. *minimální automat podmnožiny L je konečný,*
7. *kongruence $\sim_L$ má konečný index,*
8. *levá kongruence $\sim_L^\ell$ má konečný index,*
9. *pravá kongruence $\sim_L^r$ má konečný index,*
10. *množina $\{y^{-1} \cdot L \cdot z^{-1} \mid y, z \in S\}$ je konečná.*

Cvičení 1.22. Dokažte ekvivalenci podmínek 7 a 10 v předchozí větě. Všimněte si, že podmínka 10 je vlastně naruby otočená podmínka 7, neboť místo toho, jaké kontexty má daný prvek, sleduje, jaké prvky se vyskytují v daném kontextu.

1.3 Racionální podmnožiny

Dalším typem konečně popsatelných podmnožin pologrup, kterým se budeme nyní zabývat, jsou podmnožiny racionální. Zatímco rozpoznatelné podmnožiny jsou zadány konečným mechanismem, pomocí kterého lze rozhodovat příslušnost prvků do podmnožiny, racionální podmnožiny jsou vytvořeny konečným postupem z konečných podmnožin. Základní mechanismy, které se používají k popisu racionálních množin, jsou racionální výrazy a konečné nedeterministické automaty.

Racionální výrazy popisují tvorbu racionálních podmnožin z jednoprvkových podmnožin pomocí operací sjednocení, násobení po prvcích a generování podpologrupy. Formálně jsou jejich syntax a sémantika definovány následovně:

Definice 1.23. *Racionálním výrazem* nad pologrupou S rozumíme výraz, který lze vytvořit ze symbolů x, kde $x \in S$, a $\emptyset$ pomocí konečně mnoha aplikací následujících pravidel:

1. jsou-li R_1 a R_2 racionální výrazy, je i $(R_1 \cup R_2)$ racionální výraz (někdy se rovněž používá operační symbol „+“ nebo „|“),
2. jsou-li R_1 a R_2 racionální výrazy, je i $(R_1 \cdot R_2)$ racionální výraz,
3. je-li R racionální výraz, je i (R^+) racionální výraz.

Definice 1.24. Podmnožinu $\iota(R)$ pologrupy S popsanou racionálním výrazem R nad S definujeme induktivně:

0. výraz $\emptyset$ popisuje prázdnou množinu, tj. $\iota(\emptyset) = \emptyset$, a výraz x popisuje jednoprvkovou množinu $\{x\}$, tj. $\iota(x) = \{x\}$,
1. množina $\iota(R_1 \cup R_2)$ je definována jako sjednocení množin $\iota(R_1)$ a $\iota(R_2)$,
2. množina $\iota(R_1 \cdot R_2)$ je tvořena všemi součiny nějakého prvku z $\iota(R_1)$ a nějakého prvku z $\iota(R_2)$, tedy
$$\iota(R_1 \cdot R_2) = \{x \cdot y \mid x \in \iota(R_1),\ y \in \iota(R_2)\},$$
3. množina $\iota(R^+)$ je definována jako podpologrupa S generovaná množinou $\iota(R)$, tedy
$$\iota(R^+) = \{x_1 \cdots x_n \mid n \in \mathbb{N},\ x_1, \ldots, x_n \in \iota(R)\}$$
(této operaci se obvykle říká *pozitivní Kleeneho iterace*).

Říkáme, že podmnožina L pologrupy S je *racionální*, jestliže je popsatelná nějakým racionálním výrazem. Množinu všech racionálních podmnožin pologrupy S značíme $\mathrm{Rat}(S)$.

Poznámka 1.25. Při používání racionálních výrazů se obvykle nerozlišuje mezi výrazem a množinou jím popsanou, takže se dále se žádným zobrazením ι nesetkáme.

V případě, že S je monoid, používá se obvykle místo generování podpologrupy operace generování podmonoidu $L^* = L^+ \cup \{1\}$, nazývaná *Kleeneho iterace* nebo také Kleeneho hvězdička.

Cvičení 1.26. Zdůvodněte, že je-li S monoid, je možné třídu $\mathrm{Rat}(S)$ ekvivalentně definovat použitím Kleeneho iterace místo pozitivní Kleeneho iterace.

Operace použité v definici racionálních množin se nazývají *racionální operace*.

Definice 1.27. Nechť S je pologrupa. *Nedeterministickým automatem* nad S rozumíme uspořádanou pětici $\mathscr{A} = (S, Q, E, I, F)$, kde Q je nějaká množina stavů, $E \subseteq Q \times S \times Q$ je množina hran ohodnocených prvky S, podmnožina $I \subseteq Q$ je množina počátečních stavů a podmnožina $F \subseteq Q$ je množina koncových stavů. Říkáme, že automat $\mathscr{A}$ je *konečný*, jestliže množina stavů Q i množina hran E jsou konečné.

Podmnožinu

$$L(\mathscr{A}) = \{x_1 \cdots x_n \mid n \in \mathbb{N},\ \exists q_0, \ldots, q_n \in Q\colon\ q_0 \in I, q_n \in F, (q_{i-1}, x_i, q_i) \in E \text{ pro } i = 1, \ldots, n\}$$

pologrupy S složenou ze všech prvků vzniklých vynásobením ohodnocení hran na nějaké cestě z počátečního do koncového stavu nazýváme množinou *přijímanou* automatem $\mathscr{A}$.

Tvrzení 1.28 (Michael O. Rabin a Dana S. Scott, 1959)**.** *Podmnožina L pologrupy S je racionální právě tehdy, když je přijímána nějakým konečným nedeterministickým automatem nad S.*

Důkaz. Důkaz tohoto tvrzení se provede stejně jako v případě volných monoidů, přičemž tento speciální případ je možné nalézt ve většině úvodních textů do teorie formálních jazyků.

K danému racionálnímu výrazu není těžké induktivně zkonstruovat konečný nedeterministický automat, který rozpoznává jazyk popsaný tímto výrazem (racionální výraz můžeme v podstatě považovat za popis konstrukce jistého nedeterministického automatu).

Naopak, chceme-li z nedeterministického automatu vytvořit racionální výraz, použijeme k tomu zobecněné nedeterministické automaty, jejichž hrany mohou být ohodnoceny libovolným racionálním výrazem. Můžeme například nejprve upravit automat, aby obsahoval jen jeden počáteční a jeden koncový stav, a potom postupně odebírat jeho stavy a nahrazovat je hranami ohodnocenými složitějšími výrazy, které vyjadřují možné průchody přes odebíraný stav (například použijeme Kleeneho iteraci na vygenerování podpologrupy sjednocením ohodnocení všech smyček v odebíraném stavu). Takto postupujeme až do okamžiku, kdy má automat pouze počáteční a koncový stav, a pro tento automat již snadno nalezneme ekvivalentní racionální výraz. □

1.4 Vztahy mezi rozpoznatelnými a racionálními podmnožinami

Nejprve si uvědomme, jaký je v principu rozdíl mezi definicemi rozpoznatelných a racionálních podmnožin.

V případě rozpoznatelné podmnožiny, ať je její prvek zadán jakýmkoli součinem prvků pologrupy, umí jej náš automat rozpoznat. Přitom tento automat samozřejmě smí, a dokonce musí, umět pracovat se všemi prvky pologrupy. Například 0 nemůže být rozpoznatelná v grupě $\mathbb{Z}$, neboť ji lze získat součty tvaru $1+\cdots+1-1-\cdots-1$, kde potřebujeme nekonečnou paměť, abychom při čtení zleva porovnali, zda se počty čísel 1 a -1 shodují.

V případě racionální podmnožiny zase existuje pro každý její prvek nějaký zápis pomocí pevně zvolených konečně mnoha prvků, který náš automat pozná.

Obecně v pologrupách neplatí mezi racionálními a rozpoznatelnými podmnožinami ani jedna inkluze. Jak jsme viděli ve cvičení 1.4.1, příkladem racionálních podmnožin, které nejsou rozpoznatelné, jsou neprázdné konečné podmnožiny grupy $\mathbb{Z}$. Na druhou stranu si všimněte, že každá racionální podmnožina je zkonstruovaná v konečně mnoha krocích z konečně mnoha prvků, a tedy je podmnožinou nějaké konečně generované podpologrupy. Jak ukazuje následující věta, je neexistence konečné generující množiny jedinou překážkou k tomu, aby každá rozpoznatelná podmnožina dané pologrupy byla racionální.

Tvrzení 1.29 (J. D. McKnight, Jr., 1964)**.** *Pro libovolnou pologrupu S platí* $\mathrm{Rec}(S) \subseteq \mathrm{Rat}(S)$ *právě tehdy, když S je konečně generovaná.*

Důkaz. Jestliže S není konečně generovaná, potom celá S je rozpoznatelná podmnožina S, která není racionální.

Naopak, je-li S konečně generovaná, stačí pro daný konečný deterministický automat $\mathscr{A}$ uvážit konečný nedeterministický automat, jehož hrany odpovídají chování zvolených konečně mnoha generátorů pologrupy S v automatu $\mathscr{A}$ a dokázat, že vzniklý automat přijímá stejnou podmnožinu jako $\mathscr{A}$. Jelikož každý prvek x pologrupy S je součinem jejích generátorů $x = x_1 \cdots x_n$, je-li přijímán automatem $\mathscr{A}$, potom jediná cesta v novém automatu z počátečního stavu, jejíž hrany jsou ohodnoceny postupně generátory $x_1, \ldots, x_n$, vede do koncového stavu. □

1.5 Uzávěrové vlastnosti

Nejdůležitějšími uzávěrovými vlastnostmi rozpoznatelných jazyků, které budou hrát klíčovou roli v kapitole 3, jsou uzavřenost na Booleovské operace, vzory v homomorfismech a kvocienty.

Cvičení 1.30. Dokažte, že jsou-li K a L rozpoznatelné podmnožiny pologrupy S, potom $K \cup L$, $K \cap L$ a $S \setminus L$ jsou rovněž rozpoznatelné podmnožiny S. Použijte přitom definici rozpoznatelnosti pomocí homomorfismů a všimněte si, jaké algebraické konstrukce jste v důkazu použili.

Cvičení 1.31. Nechť L je rozpoznatelná podmnožina pologrupy S a $\psi\colon R \to S$ je homomorfismus pologrup. Dokažte, že potom je $\psi^{-1}(L)$ rozpoznatelná podmnožina pologrupy R. Opět přitom použijte definici rozpoznatelnosti pomocí pologrup a uvědomte si, jaký je vztah mezi syntaktickými pologrupami podmnožiny $L \subseteq S$ a podmnožiny $\psi^{-1}(L) \subseteq R$ (připomeňte si lemma 1.16).

Cvičení 1.32. Uvažujme rozpoznatelnou podmnožinu $L = (ab)^* = \{(ab)^n \mid n \in \mathbb{N}_0\}$ volného monoidu $\{a,b\}^*$ a homomorfismus $\psi\colon \{a,b\}^* \twoheadrightarrow (\mathbb{N}_0,+)^2$ daný předpisem $\psi(a) = (1,0)$ a $\psi(b) = (0,1)$. Dokažte, že podmnožina $\psi(L)$ není rozpoznatelná v $\mathbb{N}_0^2$. Dále dejte příklad rozpoznatelné podmnožiny v pologrupě $(\mathbb{N},+)$, jejíž obraz ve vložení $\mathbb{N} \hookrightarrow \mathbb{Z}$ není rozpoznatelný. Rozpoznatelné podmnožiny tedy nejsou obecně uzavřené na obrazy v surjektivních ani v injektivních homomorfismech.

Pro libovolné podmnožiny K a L pologrupy S definujeme *levý a pravý kvocient* L podle K předpisy

$$K^{-1}L = \{x \in S \mid \exists y \in K\colon yx \in L\}$$
$$LK^{-1} = \{x \in S \mid \exists y \in K\colon xy \in L\}$$

Cvičení 1.33. Dokažte, že je-li L rozpoznatelná podmnožina pologrupy S a $K \subseteq S$ libovolná podmnožina, potom jsou $K^{-1}L$ i LK^{-1} rozpoznatelné podmnožiny S. Opět si všimněte, jakou pologrupou jsou kvocienty rozpoznávány.

Nakonec se podívejme, jak je to s uzavřeností rozpoznatelných podmnožin na zbylé racionální operace. Jak uvidíme dále, v pologrupách, kde se rozpoznatelné podmnožiny běžně studují, je součin libovolných rozpoznatelných podmnožin opět rozpoznatelná podmnožina. Proto k prokázání, že obecně třída rozpoznatelných podmnožin na součin uzavřená není, potřebujeme zkonstruovat nějakou méně obvyklou pologrupu.

Cvičení 1.34 (Shmuel Winograd)**.** Uvažme komutativní pologrupu S, která vznikne z grupy $(\mathbb{Z},+)$ přidáním prvku a splňujícího $a+a=0$ a $a+z=z$ pro libovolné $z \in \mathbb{Z}$. Ukažte, že podmnožina $\{a\}$ je rozpoznatelná v S, ale podmnožina $\{a\}+\{a\}$ v S rozpoznatelná není.

Naopak, na iterace nejsou rozpoznatelné jazyky uzavřené ani v některých základních pologrupách, typicky v komutativních.

Cvičení 1.35. Uvažme volný komutativní monoid nad dvěma generátory $\mathbb{N}_0 \times \mathbb{N}_0$. Ukažte, že jeho jednoprvková podmnožina $\{(1,1)\}$ je rozpoznatelná, ale její pozitivní iterace $\{(n,n) \mid n \in \mathbb{N}\}$ rozpoznatelná není.

Nyní se podívejme, na které operace je uzavřená třída všech racionálních podmnožin. Nejprve si uvědomme, že je přímo z definice uzavřená na racionální operace. Jak je tomu s ostatními Booleovskými operacemi si ukážeme v následujícím cvičení.

Cvičení 1.36. Uvažme přímý součin dvou volných monoidů $M = \{a,b\}^* \times \{c\}^*$ a jeho dvě racionální podmnožiny $K = (a,c)^* \cdot (b,1)^*$ a $L = (a,1)^* \cdot (b,c)^*$. Ukažte, že $K \cap L$ není racionální podmnožina monoidu M. Zdůvodněte, že třída racionálních jazyků není obecně uzavřená ani na komplementy.

Cvičení 1.37. Ukažte, že třída racionálních podmnožin není obecně uzavřená na kvocienty ani jednoprvkovými podmnožinami, tedy pro racionální podmnožinu L a prvek x nemusí být podmnožina $\{x\}^{-1}L$ racionální.

Jelikož každý homomorfismus respektuje všechny racionální operace (tedy například $\varphi(K \cdot L) = \varphi(K) \cdot \varphi(L)$), obraz racionálního jazyka v homomorfismu je vždy racionální. Na rozdíl od rozpoznatelných jazyků ale nejsou racionální jazyky obecně uzavřené na vzory v homomorfismech.

Cvičení 1.38. Dejte příklad homomorfismu konečně generovaných pologrup $\psi\colon R \to S$ a racionální podmnožiny $L \subseteq S$, jejíž vzor $\psi^{-1}(L)$ racionální není.

Tvrzení 1.39. *Nechť $\psi\colon R \twoheadrightarrow S$ je surjektivní homomorfismus pologrup a nechť L je racionální podmnožina S. Potom existuje racionální podmnožina $K \subseteq R$ taková, že $\psi(K) = L$.*

Cvičení 1.40. Dokažte předchozí tvrzení. Dejte příklad homomorfismu $\psi\colon R \to S$ a podmnožiny $L \subseteq \psi(R)$ takové, že žádná racionální podmnožina $K \subseteq R$ splňující $\psi(K) = L$ neexistuje.

Následující tabulka shrnuje, na které operace jsou uzavřené třídy rozpoznatelných a racionálních jazyků na obecných pologrupách.

	$_\cup_$	$_\cap_$	$S\setminus_$	$_\cdot_$	$_^+$	$K^{-1}\cdot_$	$\psi(_)$	$\psi^{-1}(_)$
rozpoznatelné podmnožiny	✓	✓	✓	✗	✗	✓	✗	✓
racionální podmnožiny	✓	✗	✗	✓	✓	✗	✓	✗

1.6 Speciální případy

V této části si ukážeme, jak vypadají rozpoznatelné a racionální podmnožiny v nejčastěji studovaných typech pologrup.

Cvičení 1.41. Charakterizujte rozpoznatelné a racionální podmnožiny konečných pologrup.

1.6.1 Monoidy slov

Jedno z nejznámějších tvrzení o konečných automatech říká, že v konečně generovaných volných monoidech mají deterministické a nedeterministické konečné automaty stejnou vyjadřovací sílu, tedy pojmy rozpoznatelného a racionálního jazyka splývají. Často se pro tyto jazyky užívá název *regulární*.

Věta 1.42 (Stephen Cole Kleene, 1956)**.** *Je-li A konečná množina, potom libovolná podmnožina volného monoidu A^* je rozpoznatelná právě tehdy, když je racionální.*

Důkaz. Každý rozpoznatelný jazyk nad A je racionální podle tvrzení 1.29.

Obvyklou metodou důkazu obrácené implikace, se kterou se lze setkat ve většině úvodních kurzů formálních jazyků, je konstrukce deterministického automatu, jehož stavy jsou množiny stavů daného nedeterministického automatu. V tomto textu si ukážeme jiný důkaz, založený na přímé konstrukci pologrupy rozpoznávající racionální jazyk zadaný nedeterministickým automatem s hranami ohodnocenými písmeny (každý nedeterministický automat lze do tohoto tvaru snadno upravit). Tento důkaz mimo jiné ukazuje, že počet prvků syntaktické pologrupy jazyka přijímaného nedeterministickým automatem o n stavech je shora omezen 2^{n^2}, přestože jak při přechodu z nedeterministického na deterministický automat, tak při přechodu z deterministického automatu na syntaktickou pologrupu je v některých případech alespoň exponenciální nárůst počtu stavů nevyhnutelný.

Předpokládejme tedy, že jazyk L je přijímaný nedeterministickým automatem $\mathscr{A} = (A^*, Q, E, I, F)$, jehož všechny hrany jsou ohodnoceny písmeny, tj. $E \subseteq Q \times A \times Q$. Uvažme antihomomorfismus φ do pologrupy všech binárních relací na množině stavů Q daný předpisem $\varphi(a) = \{(p,q) \in Q \times Q \mid (p,a,q) \in E\}$ pro všechna $a \in A$. Potom pro libovolné $a_1, \dots, a_n \in A$ platí $(p,q) \in \varphi(a_1 \dots a_n) = \varphi(a_n) \circ \dots \circ \varphi(a_1)$ právě tehdy, když v automatu $\mathscr{A}$ existuje cesta z p do q ohodnocená slovem $a_1 \dots a_n$. Proto tento antihomomorfismus rozpoznává L množinou relací $\{R \subseteq Q \times Q \mid \exists p \in I,\ q \in F\colon (p,q) \in R\}$. Všimněte si, že obecně pro libovolnou pologrupu S takový antihomomorfismus definovat nelze, neboť relace odpovídající cestám ohodnoceným prvkem $x \cdot y$ může být větší než složení relací odpovídajících cestám ohodnoceným prvky x a y. □

Uvědomte si, že regulární jazyky mají všechny uzávěrové vlastnosti, které platí buď pro racionální nebo pro rozpoznatelné podmnožiny.

Cvičení 1.43. Nechť $\varphi\colon A^* \to T$ a $\psi\colon A^* \to U$ jsou homomorfismy rozpoznávající jazyky K a L. Ukažte, že jazyk $K \cdot L$ je rozpoznávaný homomorfismem

$$\rho\colon A^* \to \mathrm{Mat}_2(\mathscr{P}(T \times U))$$

do monoidu všech matic řádu 2 (s operací násobení matic) nad idempotentním polookruhem $(\mathscr{P}(T \times U), \cup, \cdot)$, který je daný pro všechna $a \in A$ předpisem

$$\rho(a) = \begin{pmatrix} \{(\varphi(a),1)\} & \{(\varphi(a),1),(1,\psi(a))\} \\ \emptyset & \{(1,\psi(a))\} \end{pmatrix}.$$

Dále ukažte, že jazyk L^+ je rozpoznávaný homomorfismem

$$\sigma\colon A^* \to \mathscr{P}(\mathrm{Mat}_3(\mathscr{P}(U)))$$

daným pro všechna slova $w \in A^+$ předpisem

$$\sigma(w) = \left\{ \begin{pmatrix} \{1\} & \{\psi(v)\} & W \\ \emptyset & \emptyset & \{\psi(u)\} \\ \emptyset & \emptyset & \{1\} \end{pmatrix} \middle| \begin{array}{c} n \in \mathbb{N}_0,\ u, w_1, \ldots, w_n, v \in A^*, \\ uw_1 \cdots w_n v = w, \\ W = \{\psi(w_1), \ldots, \psi(w_n)\} \end{array} \right\} \cup \left\{ \begin{pmatrix} \{1\} & \emptyset & \emptyset \\ \emptyset & \{\psi(w)\} & \emptyset \\ \emptyset & \emptyset & \{1\} \end{pmatrix} \right\}.$$

Cvičení 1.44. Spočtěte minimální automat a syntaktický monoid jazyků $\{a,b\}^*aba$ a $(ab)^*$ nad abecedou $\{a,b\}$.

Tvrzení 1.45. *Racionální podmnožiny v libovolné pologrupě S jsou právě obrazy regulárních jazyků v homomorfismech z volných konečně generovaných pologrup. Rozpoznatelné podmnožiny v libovolné konečně generované pologrupě S jsou právě ty podmnožiny, jejichž vzor v každém (ekvivalentně, nějakém surjektivním) homomorfismu z volné konečně generované pologrupy je regulární.*

Cvičení 1.46. Pomocí již dokázaných faktů o rozpoznatelných a racionálních podmnožinách dokažte toto tvrzení.

Cvičení 1.47. Použijte tvrzení 1.45 k důkazu, že průnikem racionální a rozpoznatelné podmnožiny konečně generované pologrupy je vždy racionální podmnožina.

1.6.2 Rozpoznatelné a racionální relace

Nyní se podíváme, jak vypadají rozpoznatelné podmnožiny v součinu dvou volných monoidů. Jejich charakterizace vyplývá přímo z obecného tvrzení charakterizujícího rozpoznatelné podmnožiny součinu dvou monoidů.

Tvrzení 1.48 (Mezei, nepublikováno)**.** *Podmnožina přímého součinu monoidů $M \times N$ je rozpoznatelná právě tehdy, když je konečným sjednocením množin tvaru $K \times L$, kde K je rozpoznatelná podmnožina M a L je rozpoznatelná podmnožina N.*

Cvičení 1.49. Dokažte toto tvrzení. Přitom využijte faktu, že monoidy M a N jsou přirozeně izomorfní podmonoidům monoidu $M \times N$ a tedy každý homomorfismus definovaný na $M \times N$ definuje homomorfismus na M i na N.

Ukažte, že pokud N je pologrupa, která není monoid, tak tvrzení platit nemusí, například pro $M = (\mathbb{Z}, +)$ a $N = (\mathbb{N}, +)$.

Poznámka 1.50. Toto tvrzení lze snadno rozšířit na libovolný součin konečně mnoha monoidů.

Všimněte si, že z tvrzení 1.48 ihned vyplývá, že součin dvou rozpoznatelných relací, tedy podmnožin monoidu $A^* \times B^*$, je rozpoznatelná relace.

Racionální relace (rational transductions), tedy racionální podmnožiny monoidu $A^* \times B^*$, patří k nejdůležitějším nástrojům v teorii regulárních jazyků. Nedeterministické automaty přijímající relace se obvykle nazývají *převodníky* (transducers). Racionálním relacím je věnována například klasická kniha Jeana Berstela [1]. Zmiňme tu jen několik nejdůležitějších vlastností těchto relací.

Tvrzení 1.51. *Obraz regulárního jazyka v racionální relaci je regulární. Obraz bezkontextového jazyka v racionální relaci je bezkontextový.*

Tvrzení 1.52 (Calvin C. Elgot a Jorge E. Mezei, 1965)**.** *Složení dvou racionálních relací je opět racionální relace.*

Cvičení 1.53. Dokažte předchozí dvě tvrzení.

Všimněte si, že inverze k racionální relaci je také vždy racionální relace. Tedy například pro každý homomorfismus $\varphi\colon A^* \to B^*$ je relace $\ker(\varphi) = \varphi^{-1} \circ \varphi \subseteq A^* \times A^*$ racionální.

Cvičení 1.54. Vyjádřete následující uzávěrové vlastnosti regulárních jazyků jako speciální případy uzavřenosti na obrazy v racionálních relacích:

1. obraz v substituci, kde za každé písmeno dosazujeme regulární jazyk;
2. průnik s pevně daným regulárním jazykem;
3. součin s pevně daným regulárním jazykem (tedy $L \mapsto L \cdot K$);
4. kvocienty podle pevně daného regulárního jazyka (tedy $L \mapsto L \cdot K^{-1}$ a $L \mapsto K^{-1} \cdot L$);
5. promíchání (shuffle) s pevně daným regulárním jazykem (tedy $L \mapsto \{u_1 v_1 \dots u_n v_n \mid n \in \mathbb{N},\ u_i, v_i \in A^*,\ u_1 \dots u_n \in L,\ v_1 \dots v_n \in K\}$);
6. faktory, prefixy, sufixy či podslova (tedy $L \mapsto \mathrm{fac}(L)$ a podobně);
7. slova s Hammingovou vzdáleností právě (méně než, více než) dané $k \in \mathbb{N}$ od slov původního jazyka (Hammingova vzdálenost slov stejné délky je počet pozic, na kterých se tato slova liší).

Všimněte si, že v mnoha případech je možné s výhodou využít uzavřenosti třídy racionálních relací na součiny.

Poznámka 1.55. Podobně lze některé základní binární operace s regulárními jazyky vyjádřit pomocí racionálních relací v monoidu $A^* \times A^* \times A^*$.

Cvičení 1.56. Ukažte, že každou racionální relaci $\rho \subseteq A^* \times B^*$ lze psát jako kompozici $\psi \circ \mathrm{id}_L \circ \varphi^{-1}$ pro nějakou abecedu C, regulární jazyk L nad C a homomorfismy $\varphi\colon C^* \to A^*$ a $\psi\colon C^* \to B^*$. Přitom za abecedu C lze brát množinu všech hran převodníku přijímajícího ρ.

1.6.3 Grupy

Přímo z definice rozpoznatelnosti vyplývá, že podmnožina L grupy G je rozpoznatelná právě tehdy, když existuje normální podgrupa H konečného indexu v G taková, že L je sjednocením tříd rozkladu G/H. Následující tvrzení ukazuje, že předpoklad normálnosti podgrupy H je ve skutečnosti nadbytečný.

Tvrzení 1.57. *Podmnožina L grupy G je rozpoznatelná právě tehdy, když existuje podgrupa H konečného indexu v G taková, že L je sjednocením tříd rozkladu $H\backslash G$.*

Poznámka 1.58. Všimněte si, že v tomto tvrzení nezáleží na tom, zda mluvíme o pravých či levých třídách rozkladu, neboť podle cvičení 1.33 je levá třída

$$gH = (g^{-1})^{-1} \cdot (Hg) \cdot g^{-1}$$

rozpoznatelná právě tehdy, když je rozpoznatelná pravá třída

$$Hg = g^{-1} \cdot (gH) \cdot (g^{-1})^{-1}.$$

Použijeme proto pravé třídy, které přímo odpovídají automatům.

Důkaz. Stačí si uvědomit, že přirozená akce grupy G na množině $H\backslash G$ zadává strukturu deterministického automatu přijímajícího L. □

Z předchozího tvrzení tedy plyne, že podgrupa je rozpoznatelná právě tehdy, když má konečný index. O trochu obtížnější je dokázat, že podgrupa v libovolné grupě je racionální právě tehdy, když je konečně generovaná (elegantní důkaz lze nalézt v [8] jako proposition II.6.2; jestli se ovšem autor tohoto textu nemýlí, je v tomto důkazu nutné změnit definice jednoduché i skoro jednoduché cesty, aby fungoval). Díky těmto charakterizacím rozpoznatelných a racionálních podgrup vidíme, že v případě grup je tvrzení 1.29 zobecněním známého faktu, že v konečně generované grupě je každá podgrupa konečného indexu konečně generovaná.

Obzvláště zajímavé vlastnosti mají racionální podmnožiny ve volných grupách, a tyto podmnožiny byly hojně studovány.

1.6.4 Volné komutativní monoidy

Protože volné komutativní monoidy nad konečně mnoha generátory jsou přímým součinem kopií monoidu $(\mathbb{N}_0,+)$, jak vypadají jejich rozpoznatelné podmnožiny nám říká tvrzení 1.48.

Následující pojem, charakterizující v těchto monoidech racionalitu, je analogií pojmu afinního podprostoru v situaci, kdy je možné provádět pouze lineární kombinace, kde všechny koeficienty jsou nezáporná celá čísla. Podmnožina monoidu $(\mathbb{N}_0,+)^n$ tvaru

$$\{(a_1,\dots,a_n)+k_1\cdot(b_{1,1},\dots,b_{1,n})+\dots+k_m\cdot(b_{m,1},\dots,b_{m,n}) \mid k_1,\dots,k_m\in\mathbb{N}_0\}$$

pro nějaká $a_1,\dots,a_n,b_{1,1},\dots,b_{m,n}\in\mathbb{N}_0$ se nazývá *lineární*. Říkáme, že podmnožina monoidu $(\mathbb{N}_0,+)^n$ je *pololineární*, jestliže je sjednocením konečně mnoha lineárních podmnožin.

Cvičení 1.59. Dokažte, že podmnožina monoidu $(\mathbb{N}_0,+)^n$ je racionální právě tehdy, když je pololineární.

Seymour Ginsburg a Edwin H. Spanier v roce 1964 objevili, že pololineární množiny jsou právě podmnožiny $(\mathbb{N}_0)^n$ definovatelné v Presburgerově aritmetice, tj. formulí logiky 1. řádu používající pouze operaci sčítání. Důsledkem této ekvivalence je například uzavřenost pololineárních množin na průnik a komplement a algoritmická rozhodnutelnost základních otázek o těchto množinách.

Poznámka 1.60. Uvažujme přirozenou projekci z volného monoidu nad n generátory $\{a_1,\dots,a_n\}^*$ na $(\mathbb{N}_0)^n$ danou předpisem $p(a_i)=(0,\dots,0,1,0,\dots,0)$, kde 1 je na i-té pozici. Jedna ze základních vlastností bezkontextových jazyků, Parikhova věta, říká, že pro každý bezkontextový jazyk $L\subseteq\{a_1,\dots,a_n\}^*$ je jeho obraz $p(L)$ pololineární množina. Protože z tvrzení 1.39 víme, že racionální podmnožiny $(\mathbb{N}_0)^n$ jsou právě obrazy regulárních jazyků v projekci p, plyne z Parikhovy věty, že každý bezkontextový jazyk je až na proházení pořadí písmen stejný jako nějaký jazyk regulární.

1.6.5 Monoidy stop*

Důležitým společným zobecněním volných monoidů a volných komutativních monoidů jsou volné částečně komutativní monoidy, obvykle nazývané monoidy stop. Praktickou motivací pro studium těchto monoidů jsou rozhodovací problémy týkající se vzájemně si konkurujících procesů. Reprezentujeme-li každý proces v systému jedním písmenem abecedy, potom procesy, u kterých nezáleží na pořadí provedení, odpovídají komutujícím písmenům, kdežto písmena odpovídající konkurujícím si procesům spolu nekomutují.

Formálně uvážíme na abecedě A libovolnou ireflexivní a symetrickou binární *relaci nezávislosti* I. Její komplement $(A\times A)\setminus I$ se nazývá *relace závislosti* a značí D. *Monoid stop* $\mathbb{M}(A,I)$ příslušný relaci I definujeme jako kvocient $A^*/{\sim_I}$, kde $\sim_I$ je kongruence

generovaná relací $\{(ab, ba) \mid a\, I\, b\}$. Prvky tohoto monoidu se zjednodušeně značí $[u]_I$, kde $u \in A^*$, a nazývají se *stopy*.

Pro rozpoznatelné podmnožiny v monoidech stop existuje charakterizace zobecňující charakterizaci rozpoznatelných jazyků pomocí racionálních operací. Jediným rozdílem oproti volným monoidům je, že povolíme aplikovat iteraci pouze na takzvané souvislé podmnožiny. Říkáme, že stopa $[u]_I$ je *souvislá*, jestliže zúžení relace D na písmena vyskytující se ve slově u tvoří souvislý graf. Podmnožina $L \subseteq \mathbb{M}(A, I)$ se nazývá *souvislá*, jestliže všechny její prvky jsou souvislé stopy. Všimněte si, že ve volném monoidu $\mathbb{M}(A, \emptyset)$ je každá podmnožina souvislá.

Věta 1.61 (Edward Ochmański, 1985)**.** *Podmnožina $L \subseteq \mathbb{M}(A, I)$ je rozpoznatelná právě tehdy, když lze získat z jednoprvkových podmnožin pomocí racionálních operací, přičemž iterace je aplikována pouze na souvislé podmnožiny.*

Příklad 1.62. Je-li $a\, I\, b$, není množina $\{[ab]_I\}$ souvislá a její iterace není rozpoznatelná.

Důkaz této věty je svým rozsahem mimo možnosti tohoto kurzu. Jeho podstatná část je založena na pojmu hodnosti množiny slov vzhledem k I, který si nyní vysvětlíme. Uvážíme-li libovolnou podmnožinu $M \subseteq A^*$, potom její uzávěr vzhledem ke kongruenci $\sim_I$ je $\overline{M} = \{u \in A^* \mid \exists v \in M\colon [u]_I = [v]_I\}$. Přitom je možné snadno ukázat, že pro libovolná slova $u, v \in A^*$ platí $uv \in \overline{M}$ právě tehdy, když existují slova $u_0, \ldots, u_n, v_0, \ldots, v_n \in A^*$ taková, že $[u_0 \ldots u_n]_I = [u]_I$, $[v_0 \ldots v_n]_I = [v]_I$, $u_0 v_0 \ldots, u_n v_n \in M$ a současně pro všechna $i < j$ jsou všechna písmena slova v_i nezávislá na všech písmenech slova u_j. *Hodností* (rank) jazyka M rozumíme (pokud existuje) nejmenší číslo $n_0 \in \mathbb{N}_0$ takové, že pro libovolnou dvojici slov $u, v \in A^*$ splňující $uv \in \overline{M}$ existuje výše uvedený rozklad s $n \leq n_0$. Dá se dokázat, že je-li M regulární jazyk konečné hodnosti, potom je jeho uzávěr $\overline{M}$ rovněž regulární. Podmínka konečné hodnosti n_0 nám totiž v podstatě říká, že po přečtení libovolného prefixu u daného slova uv si potřebujeme pamatovat pouze informace o rozkladech slova u na nejvýše n_0 faktorů (pro každý z těchto faktorů musíme znát, jaké má kontexty v M a jaká obsahuje písmena).

Díky tvrzení 1.45 nyní víme, že má-li regulární jazyk konečnou hodnost, potom množina stop jím určená je nejen racionální, ale dokonce rozpoznatelná. Abychom například dokázali, že součin rozpoznatelných množin stop K a L je opět rozpoznatelná množina, potřebujeme ověřit, že uzávěr množiny slov $\{u \in A^* \mid [u]_I \in K\} \cdot \{u \in A^* \mid [u]_I \in L\}$ je regulární jazyk. Pro libovolné jazyky $M, N \subseteq A^*$ je ovšem možné dokázat, že hodnost jazyka $\overline{M} \cdot \overline{N}$ je nejvýše 1, a je-li navíc příslušná množina stop $\{[u]_I \mid u \in M\}$ souvislá, lze také dokázat, že hodnost jazyka $(\overline{M})^*$ je nejvýše $2 \cdot |A|$. Proto jsou rozpoznatelné množiny stop uzavřené na součiny a na iterace souvislých množin.

Důkaz opačné implikace věty, tedy že každá rozpoznatelná množina stop lze zadat takto omezeným racionálním výrazem, je založen na reprezentaci stop v lexikografické normální formě.

Více informací o jazycích stop obsahuje například Handbook of formal languages [7] v osmé kapitole třetího svazku.

Kapitola 2

Struktura konečných pologrup

Abychom si mohli ukázat některé zajímavé výsledky o regulárních jazycích, které se pomocí algebraických metod podařilo získat, musíme se nejprve trochu vyznat ve struktuře konečných pologrup. Doporučeným textem pro úvod do teorie pologrup je kniha Pierra Antoina Grilleho [4]. Pinova nedokončená kniha [6] pokrývá celou tuto kapitolu včetně ukázek použití strukturní teorie konečných pologrup k důkazu výsledků o regulárních jazycích.

2.1 Podgrupy

Důležitým pojmem používaným v následujícím textu je zobecnění pojmu podgrupy z grup na podgrupy libovolné pologrupy. *Podgrupou* pologrupy S rozumíme libovolnou podpologrupu pologrupy S, která je sama grupou. Všimněte si ale, že i v případě, kdy pologrupa S obsahuje neutrální prvek, nemusí být tento prvek současně neutrálním prvkem každé její podgrupy. Neutrální prvek každé podgrupy ovšem musí být idempotentní, a naopak, každý idempotentní prvek $x \in S$ sám tvoří jednoprvkovou podgrupu $\{x\}$. Je-li S grupa, potom obsahuje jediný idempotentní prvek a inverze ke všem jejím prvkům jsou určeny jednoznačně; proto se v tomto případě právě definovaný pojem podgrupy skutečně shoduje s pojmem obvyklým v teorii grup. Prvek pologrupy, který leží v nějaké její podgrupě, se nazývá *grupový*.

2.2 Jednogenerované podpologrupy

Podobně jako tomu bývá zvykem v případě grup, si nejprve popíšeme, jak vypadají pologrupy generované jedním prvkem. Víme, že všechny jednogenerované (neboli cyklické) grupy jsou buď izomorfní $(\mathbb{Z}, +)$ nebo $(\mathbb{Z}_n, +)$ pro nějaké $n \in \mathbb{N}$.

Pologrupa $\langle x \rangle$ generovaná prvkem x sestává právě z prvků x^n pro $n \in \mathbb{N}$. Pokud jsou pro různá n všechny tyto prvky různé, potom je $\langle x \rangle$ izomorfní pologrupě $(\mathbb{N}, +)$.

Zabývejme se nyní opačným případem. Zvolme $i \in \mathbb{N}$ jako nejmenší číslo takové, že existuje $m \in \mathbb{N}$ splňující $x^{i+m} = x^i$. Toto číslo se nazývá *index* prvku x. Dále zvolme $p \in \mathbb{N}$, nazývané *perioda* prvku x, jako nejmenší číslo splňující $x^{i+p} = x^i$.

Cvičení 2.1. Dokažte, že má-li prvek x index i a periodu p, potom pologrupa $\langle x \rangle$ jím generovaná obsahuje právě $i + p - 1$ prvků. Dále ukažte, že prvek $x^{i \cdot p}$ je jediným idempotentním prvkem v této pologrupě (místo $i \cdot p$ můžeme brát libovolný násobek p větší nebo roven i) a že podpologrupa $\{x^i, \ldots, x^{i+p-1}\}$ je cyklická grupa řádu p generovaná prvkem $x^{i \cdot p+1}$.

Reprezentace této pologrupy akcí jejího generátoru tedy vypadá následovně:

$$\begin{tikzcd}
 & & & x^{i+1} \arrow[r] & \cdots \arrow[dr] & \\
x \arrow[r] & x^2 \arrow[r] & \cdots \arrow[r] & x^i = x^{i+p} \arrow[ur] & & \vdots \arrow[dl] \\
 & & & & x^{i+p-1} \arrow[ul] & \cdots \arrow[l]
\end{tikzcd}$$

Z předchozího cvičení vyplývá, že jediný idempotentní prvek v pologrupě $\langle x \rangle$ lze vždy získat jako limitu $\lim\limits_{n \to \infty} x^{n!}$ posloupnosti, která se vždy nejpozději na prvku $x^{|\langle x \rangle|!}$ stabilizuje. Tento idempotentní prvek budeme značit x^ω. Všimněte si, že v libovolné pologrupě jsou její idempotentní prvky právě prvky tvaru x^ω.

Pologrupa se nazývá *periodická*, jestliže každá její jednogenerovaná podpologrupa je konečná. Na každé periodické pologrupě máme tedy korektně definovánu unární operaci ω s následujícími vlastnostmi: $x^\omega \cdot x^\omega = x^\omega$, $x \cdot x^\omega = x^\omega \cdot x$. Pro libovolný prvek x má proto smysl psát výrazy tvaru $x^{\omega+n}$, kde $n \in \mathbb{N}$. Budeme rovněž používat zápis $x^{\omega-n}$, který označuje inverzi prvku $x^{\omega+n}$ v podgrupě generované $x^{\omega+1}$. Uvědomte si, že pro všechna $n \in \mathbb{N}$ platí

$$(x^n)^\omega = (x^{\omega+n})^\omega = (x^{\omega-n})^\omega = x^\omega.$$

Cvičení 2.2. Dokažte, že periodický monoid S je grupou právě tehdy, když pro všechny prvky $x \in S$ platí $x^\omega = 1$.

Jak uvidíme dále, mnoho důležitých vlastností konečných pologrup plyne ve skutečnosti už z jejich periodicity.

2.3 Greenovy relace

Základní nástroj používaný k popisu struktury pologrup zavedl James Alexander Green v roce 1951. Jedná se o několik relací ekvivalence, které popisují dělitelnost prvků v pologrupě, a tím také strukturu ideálů této pologrupy. Protože pologrupy nejsou obecně komutativní, musíme uvažovat zvlášť dělitelnost zleva (které odpovídá relace $\mathscr{L}$), zprava ($\mathscr{R}$)

a oboustrannou ($\mathscr{J}$). Tyto relace jsou doplněny dvěma dalšími, odvozenými, relacemi $\mathscr{H}$ a $\mathscr{D}$, které umožňují názorný grafický popis struktury pologrup. Ve většině následujícího textu budeme implicitně předpokládat, že se zabýváme jedinou pologrupou S. Pokud bychom uvažovali Greenovy relace ve více pologrupách současně, museli bychom je doplnit patřičnými indexy, například $\mathscr{L}_S$.

2.3.1 Definice a základní vlastnosti

Levým (*pravým*) *ideálem* pologrupy S rozumíme podmnožinu $I \subseteq S$ splňující $S \cdot I \subseteq I$ (respektive $I \cdot S \subseteq I$). Podmnožinu $I \subseteq S$ nazýváme *ideálem* S, je-li současně levým i pravým ideálem, tedy jestliže platí $S^1 \cdot I \cdot S^1 \subseteq I$.

Cvičení 2.3. Ukažte, že levý, pravý a oboustranný ideál generovaný podmnožinou $M \subseteq S$ jsou rovny $S^1 \cdot M$, $M \cdot S^1$ a $S^1 \cdot M \cdot S^1$.

Nejprve definujeme tři předuspořádání na pologrupě S, každé z nich dvěma ekvivalentními podmínkami, z nichž jedna mluví o vztahu mezi ideály generovanými danými prvky a druhá o jejich vzájemné dělitelnosti:

$$\begin{aligned}
y \leq_{\mathscr{L}} x \quad &\iff \quad S^1 y \subseteq S^1 x \quad \iff \quad y \in S^1 x, \\
y \leq_{\mathscr{R}} x \quad &\iff \quad y S^1 \subseteq x S^1 \quad \iff \quad y \in x S^1, \\
y \leq_{\mathscr{J}} x \quad &\iff \quad S^1 y S^1 \subseteq S^1 x S^1 \quad \iff \quad y \in S^1 x S^1.
\end{aligned}$$

Všimněte si, že nerovnosti $\leq_{\mathscr{L}}$, $\leq_{\mathscr{R}}$ a $\leq_{\mathscr{J}}$ zapisujeme v opačném směru než je obvyklé v případě relace dělitelnosti $|$, tedy že nerovnost $y \leq_{\mathscr{J}} x$ vlastně znamená, že prvek y je dělitelný prvkem x.

Protože násobení prvku z jedné strany nijak neovlivňuje jeho násobení z druhé strany, jsou předuspořádání $\leq_{\mathscr{L}}$ a $\leq_{\mathscr{R}}$ jednostranně kompatibilní s operací násobení, tedy z nerovnosti $y \leq_{\mathscr{L}} x$ plyne $yz \leq_{\mathscr{L}} xz$ pro všechna $z \in S$, a podobně z $y \leq_{\mathscr{R}} x$ plyne $zy \leq_{\mathscr{R}} zx$. Všimněte si, že generování ideálů podmnožinou $M \subseteq S$ můžeme zpětně vyjádřit pomocí těchto předuspořádání jako uzávěr směrem dolů:

$$\begin{aligned}
S^1 M &= \{y \in S \mid \exists x \in M \colon y \leq_{\mathscr{L}} x\}, \\
M S^1 &= \{y \in S \mid \exists x \in M \colon y \leq_{\mathscr{R}} x\}, \\
S^1 M S^1 &= \{y \in S \mid \exists x \in M \colon y \leq_{\mathscr{J}} x\}.
\end{aligned}$$

První tři *Greenovy ekvivalence* definujeme jako relace ekvivalence příslušející těmto předuspořádáním:

$$\begin{aligned}
x \mathscr{L} y \quad &\iff \quad x \leq_{\mathscr{L}} y \ \& \ y \leq_{\mathscr{L}} x \quad \iff \quad S^1 x = S^1 y, \\
x \mathscr{R} y \quad &\iff \quad x \leq_{\mathscr{R}} y \ \& \ y \leq_{\mathscr{R}} x \quad \iff \quad x S^1 = y S^1, \\
x \mathscr{J} y \quad &\iff \quad x \leq_{\mathscr{J}} y \ \& \ y \leq_{\mathscr{J}} x \quad \iff \quad S^1 x S^1 = S^1 y S^1.
\end{aligned}$$

Prvky tedy prohlásíme za ekvivalentní, jestliže generují stejný (levý, pravý) ideál. Třídu prvku $x \in S$ v ekvivalenci $\mathscr{J}$ budeme značit J_x, a podobně pro ostatní Greenovy ekvivalence.

Díky jednostranné kompatibilitě předuspořádání $\leq_{\mathscr{L}}$ a $\leq_{\mathscr{R}}$ s násobením je relace $\mathscr{L}$ pravá kongruence a relace $\mathscr{R}$ levá kongruence pologrupy S. Z triviálních inkluzí $\leq_{\mathscr{L}} \subseteq \leq_{\mathscr{J}}$ a $\leq_{\mathscr{R}} \subseteq \leq_{\mathscr{J}}$ dostáváme, že platí $\mathscr{L} \subseteq \mathscr{J}$ a $\mathscr{R} \subseteq \mathscr{J}$.

Původní předuspořádání nám samozřejmě indukují částečné uspořádání $\mathscr{L}$-tříd, $\mathscr{R}$-tříd a $\mathscr{J}$-tříd. Všimněte si, že pokud vynásobíme nějaký prvek x zleva (zprava, z jakékoli strany) libovolným prvkem, výsledný prvek bude patřit do stejné nebo nižší $\mathscr{L}$-třídy (respektive $\mathscr{R}$-třídy, $\mathscr{J}$-třídy) než x.

Cvičení 2.4. Dokažte, že každá pologrupa má nejvýše jednu minimální $\mathscr{J}$-třídu. Obsahuje-li pologrupa nulový prvek, pak tento prvek tvoří jednoprvkovou nejnižší $\mathscr{J}$-třídu.

Následující lemma uvádí základní vlastnost relací $\mathscr{L}$ a $\mathscr{R}$, která vyjadřuje nezávislost násobení daného prvku zleva na násobení zprava.

Lemma 2.5. $\mathscr{L} \circ \mathscr{R} = \mathscr{R} \circ \mathscr{L}$.

Důkaz. Předpokládejme, že $(x,z) \in \mathscr{L} \circ \mathscr{R}$. Potom existuje $y \in S$ splňující $x \mathscr{R} y \mathscr{L} z$. Podle definice Greenových ekvivalencí tedy máme $y = xs$, $x = yt$, $z = uy$ a $y = vz$ pro nějaká $s,t,u,v \in S^1$. Uvažujme prvek $w = ux$. Potom platí $v\underline{w} = vu\underline{x} = v\underline{uyt} = \underline{vz}t = yt = x$, a tedy $x \mathscr{L} w$. Dále můžeme spočítat $\underline{z}t = u\underline{yt} = ux = w$ a $\underline{w}s = u\underline{xs} = uy = z$, což znamená, že $w \mathscr{R} z$. Celkem tedy dostáváme požadované $(x,z) \in \mathscr{R} \circ \mathscr{L}$.

Druhá inkluze se ukáže symetricky. □

Z tohoto lemmatu ihned plyne, že kompozice relací $\mathscr{L}$ a $\mathscr{R}$ je opět relace ekvivalence, kterou značíme $\mathscr{D}$. Platí tedy, že $x \mathscr{D} y$ právě tehdy, když $R_x \cap L_y \neq \emptyset$, nebo ekvivalentně $L_x \cap R_y \neq \emptyset$. Přímo z definice plyne, že $\mathscr{D}$ je nejmenší relace ekvivalence obsahující $\mathscr{L}$ i $\mathscr{R}$. Proto musí být obsažena v $\mathscr{J}$, tedy vždy platí $\mathscr{D} \subseteq \mathscr{J}$.

Konečně, poslední Greenovou ekvivalencí je průnik $\mathscr{L} \cap \mathscr{R}$, který se značí $\mathscr{H}$.

Poznámka 2.6. V komutativní pologrupě jsou si všechny Greenovy relace rovny.

Příklad 2.7. Snadno se vidí, že ve volném monoidu A^* platí $u \leq_{\mathscr{L}} v$ právě tehdy, když slovo v je sufixem slova u. Symetricky, $u \leq_{\mathscr{R}} v$ právě tehdy, když v je prefixem u. Podobně platí $u \leq_{\mathscr{J}} v$ právě tehdy, když v je faktorem u. Dostáváme tedy

$$u \mathscr{J} v \iff u \mathscr{D} v \iff u \mathscr{L} v \iff u \mathscr{R} v \iff u \mathscr{H} v \iff u = v.$$

Pologrupám s touto vlastností se říká *$\mathscr{J}$-triviální*.

Díky komutativitě relací $\mathscr{L}$ a $\mathscr{R}$ je možné každou $\mathscr{D}$-třídu názorně reprezentovat graficky jako tabulku, kde řádky představují $\mathscr{R}$-třídy, sloupce $\mathscr{L}$-třídy a jednotlivé buňky $\mathscr{H}$-třídy. Například situaci z předchozího důkazu můžeme znázornit následovně:

$$\begin{array}{ccc} & \xrightarrow{\cdot s} & \\ & \begin{array}{|c|c|c|} \hline x & & y \\ \hline & & \\ \hline w & & z \\ \hline \end{array} & \\ & \xleftarrow[\cdot t]{} & \end{array}$$

$u\cdot \downarrow$ $\uparrow v\cdot$

V této tabulce jsou znázorněny dvě $\mathscr{R}$-třídy a dvě $\mathscr{L}$-třídy, které patří do stejné $\mathscr{D}$-třídy a protínají se ve čtyřech $\mathscr{H}$-třídách reprezentovaných prvky $x, y, w, z \in S$.

Všimněte si, že v důkazu jsme ověřili, že násobení prvky s a t zprava přenáší prvky $\mathscr{D}$-třídy mezi příslušnými sloupci v obou uvažovaných řádcích. Podobně funguje násobení prvky u a v zleva pro řádky. Toto násobení ve skutečnosti realizuje bijekce mezi všemi $\mathscr{H}$-třídami v rámci jedné $\mathscr{D}$-třídy.

Lemma 2.8 (James Alexander Green, 1951)**.** *Nechť $(x,y) \in \mathscr{R}$ a nechť $s,t \in S^1$ splňují $xs = y$ a $yt = x$. Potom předpisy $w \mapsto ws$ a $z \mapsto zt$ definují vzájemně inverzní bijekce mezi $\mathscr{L}$-třídami L_x a L_y, které zachovávají $\mathscr{R}$-třídy.*

Symetricky, pokud platí $(y,z) \in \mathscr{L}$ a prvky $u,v \in S^1$ splňují $uy = z$ a $vz = y$, potom předpisy $x \mapsto ux$ a $w \mapsto vw$ definují vzájemně inverzní bijekce mezi $\mathscr{R}$-třídami R_y a R_z, které zachovávají $\mathscr{L}$-třídy.

Cvičení 2.9. Použitím podobných argumentů jako v důkazu lemmatu 2.5 dokažte toto tvrzení.

Důsledek 2.10. *Všechny $\mathscr{H}$-třídy obsažené ve stejné $\mathscr{D}$-třídě mají stejnou mohutnost.*

Nechť nyní H je libovolná $\mathscr{H}$-třída. Uvažujme všechna zobrazení $f_s \colon H \to H$, kde $f_s(x) = xs$, určená násobením zprava nějakým prvkem $s \in S^1$ takovým, že pro všechna $x \in H$ patří prvek xs opět do H. Množina těchto zobrazení je zřejmě uzavřená na kompozice, přičemž platí $f_t \circ f_s = f_{st}$. Díky lemmatu 2.8 víme, že k tomu, aby xs náleželo do H pro všechna $x \in H$, stačí, aby tomu tak bylo pro jediné $x \in H$. Přitom z hodnoty xs pro jedno $x \in H$ vyplývá, jaká je hodnota ys pro všechna $y \in H$: každé $y \in H$ lze totiž psát ve tvaru $y = ux$ pro nějaké $u \in S^1$, a tedy $ys = u \cdot (xs)$. Dále z lemmatu 2.8 plyne, že všechna zobrazení f_s jsou bijekce a inverzi každého z nich je rovněž možné zadat jako násobení nějakým prvkem S^1. Dostáváme tedy grupu $\Gamma(H) = \{f_s \mid s \in S^1,\ Hs = H\}$ vzhledem ke skládání zobrazení. Tato grupa se nazývá pravá *Schützenbergerho grupa* $\mathscr{H}$-třídy H. Všimněte si, že díky lemmatu 2.8 můžeme místo o pravé Schützenbergerho grupě $\mathscr{H}$-třídy mluvit o Schützenbergerho grupě příslušné $\mathscr{L}$-třídy.

Pomocí násobení zleva můžeme analogicky definovat *levou Schützenbergerho grupu* dané $\mathscr{H}$-třídy, případně $\mathscr{R}$-třídy.

Lemma 2.11. *Levá a pravá Schützenbergerho grupa třídy H jsou izomorfní a mají stejnou mohutnost jako H.*

Lemma 2.12. *Schützenbergerho grupy libovolných $\mathscr{H}$-tříd ve stejné $\mathscr{D}$-třídě jsou izomorfní.*

Cvičení 2.13. Dokažte předchozí lemmata. Přitom je užitečné si uvědomit, že stačí sledovat chování každého prvku Schützenbergerho grupy pouze na jednom pevně zvoleném prvku H.

Nyní si ukážeme, že pokud výsledkem vynásobení nějakých dvou prvků jisté $\mathscr{H}$-třídy pologrupy S je opět prvek této $\mathscr{H}$-třídy, potom je tato $\mathscr{H}$-třída podgrupou S. K důkazu tohoto tvrzení použijeme následující kritérium pro rozpoznání, zda daná pologrupa je grupa.

Lemma 2.14. *Pro libovolnou pologrupu T jsou následující podmínky ekvivalentní:*

1. *Pologrupa T má právě jednu $\mathscr{H}$-třídu.*
2. *Pologrupa T je neprázdná a pro všechny prvky $x, y \in T$ existují prvky $s, t \in T$ takové, že $xs = tx = y$.*
3. *Pologrupa T je grupa.*

Jestliže v neprázdné pologrupě T pro všechny prvky $x, y \in T$ existují prvky $s, t \in T$ takové, že $xs = tx = y$, potom T je grupa.

Cvičení 2.15. Dokažte toto lemma. Při důkazu lze s výhodou využít faktu, že levý a pravý neutrální prvek si musejí být rovny, a analogického tvrzení pro levé a pravé inverze.

Lemma 2.16 (James Alexander Green, 1951)**.** *Pro každou $\mathscr{H}$-třídu H buď platí $H^2 \cap H = \emptyset$, nebo je H podgrupa izomorfní své Schützenbergerho grupě.*

Důkaz. Pokud existuje prvek $x \in H^2 \cap H$, máme $x = yz$ pro jistá $y, z \in H$. Potom podle lemmatu 2.8 definují násobení zleva prvkem y a násobení zprava prvkem z bijekce na H, a tedy platí $yH = H$ a $Hz = H$. Dalším použitím lemmatu 2.8 se nyní dozvíme, že násobení zprava i zleva libovolným prvkem H definuje bijekci na H. Jinými slovy, pro všechna $h \in H$ platí $Hh = H = hH$. Třída H je tedy podpologrupou a navíc splňuje předpoklady lemmatu 2.14, a proto je ve skutečnosti podgrupou. Nyní si stačí všimnout, že každá bijekce patřící do Schützenbergerho grupy $\mathscr{H}$-třídy H je v tomto případě realizována nějakým prvkem H, neboť pro všechna $h \in H$ platí $h = h \cdot e$, kde e je neutrální prvek grupy H. Přitom kompozice těchto bijekcí odpovídá v případě levé Schützenbergerho grupy násobení prvků, kterými jsou tyto bijekce realizovány, a proto je H izomorfní své Schützenbergerho grupě. $\square$

Lemma 2.17. *Každá podgrupa pologrupy S je obsažená v nějaké podgrupě maximální vzhledem k inkluzi. Přitom maximální podgrupy jsou právě $\mathscr{H}$-třídy, které obsahují idempotentní prvek. Proto je každý grupový prvek x obsažen v právě jedné maximální podgrupě, která obsahuje všechny podgrupy obsahující x.*

Cvičení 2.18. Dokažte toto lemma.

Cvičení 2.19. Pro libovolnou idempotentní pologrupu S dokažte postupně následující tvrzení:

1. Pro libovolné prvky $x, y, z \in S$ platí $xyzy \mathscr{R} xyz$.

2. Jsou-li $u = x_1 \cdots x_m$ a $w = y_1 \cdots y_n$ součiny prvků pologrupy S takové, že každý z prvků y_i je roven jednomu z prvků x_j, potom platí $uw \mathscr{R} u$.

3. Pologrupa S splňuje každou identitu $uvw = uw$, kde u a w jsou slova obsahující právě tytéž proměnné a slovo v obsahuje pouze proměnné obsažené v u a w.

Cvičení 2.20. Dokažte, že je-li S periodická pologrupa, potom pro každý prvek $x \in S$ a každý idempotentní prvek $e \in S$ platí $x \mathscr{H} e$ právě tehdy, když $x^\omega = e$ a $xe = x$.

Cvičení 2.21. Dokažte, že pro libovolnou množinu Q jsou Greenovy relace v úplném transformačním monoidu $\mathscr{T}(Q)$ popsatelné následovně:

$$\begin{aligned}
\rho \leq_{\mathscr{L}} \sigma \quad &\iff \quad \ker(\rho) \supseteq \ker(\sigma), \\
\rho \leq_{\mathscr{R}} \sigma \quad &\iff \quad \mathrm{Im}(\rho) \subseteq \mathrm{Im}(\sigma), \\
\rho \leq_{\mathscr{J}} \sigma \quad &\iff \quad |\mathrm{Im}(\rho)| \leq |\mathrm{Im}(\sigma)|, \\
\rho \mathscr{L} \sigma \quad &\iff \quad \ker(\rho) = \ker(\sigma), \\
\rho \mathscr{R} \sigma \quad &\iff \quad \mathrm{Im}(\rho) = \mathrm{Im}(\sigma), \\
\rho \mathscr{J} \sigma \quad &\iff \quad \rho \mathscr{D} \sigma \quad \iff \quad |\mathrm{Im}(\rho)| = |\mathrm{Im}(\sigma)|.
\end{aligned}$$

Všimněte si, že v případě, kdy je množina Q nespočetná, musíme k důkazu použít axiom výběru. V případech, kdy Q je konečná nebo spočetná množina, se pokuste určit, jak jsou $\mathscr{J}$-třídy $\mathscr{T}(Q)$ uspořádány, kolik obsahují $\mathscr{L}$-tříd a $\mathscr{R}$-tříd, jaké jsou mohutnosti jejich $\mathscr{H}$-tříd, které $\mathscr{H}$-třídy jsou podgrupy, případně jakým grupám jsou tyto $\mathscr{H}$-třídy izomorfní.

Cvičení 2.22. Podobně jako v předchozím příkladu popište Greenovy relace na monoidu všech injektivních transformací množiny $\mathbb{N}$.

2.3.2 Greenovy relace v konečných pologrupách

Většina základních vlastností Greenových relací platných v konečných pologrupách platí ve skutečnosti již díky jejich periodicitě, přičemž důkazy těchto vlastností jsou založeny na faktu, že pro každý prvek existuje nějaká jeho mocnina, která je idempotentní.

Lemma 2.23. *Nechť S je periodická pologrupa, $x \in S$ a $s,t \in S^1$. Pokud platí $x \mathscr{J} sxt$, potom platí $sx \mathscr{L} x \mathscr{R} xt$.*

Důkaz. Podle předpokladu existují $u,v \in S^1$ takové, že $usxtv = x$. Z tohoto vztahu se snadno indukcí ukáže rovnost $x = (us)^n x(tv)^n$ pro všechna $n \in \mathbb{N}$. Díky periodicitě pologrupy S existují přirozená čísla k a ℓ, pro něž jsou prvky $(us)^k$ a $(tv)^\ell$ idempotentní. Dostáváme tedy rovnost

$$x = (us)^{k\ell} x(tv)^{k\ell} = (us)^\omega x(tv)^\omega$$

a z ní odvodíme

$$x = (us)^\omega x(tv)^\omega = (us)^\omega (us)^\omega x(tv)^\omega = (us)^\omega x = (us)^{\omega-1} u \cdot sx,$$

což dokazuje, že $sx \mathscr{L} x$. Symetrickým postupem lze ukázat, že $x \mathscr{R} xt$. □

Poznámka 2.24. Další významnou třídou pologrup, kde toto lemma (a tedy všechny jeho následující důsledky) platí jsou kompaktní topologické pologrupy.

Aplikací předchozího lemmatu získáme důkaz nejdůležitější vlastnosti konečných pologrup.

Věta 2.25 (James Alexander Green, 1951)**.** *V každé periodické pologrupě S platí $\mathscr{D} = \mathscr{J}$.*

Důkaz. Vezměme libovolné prvky $x,y \in S$ splňující $x \mathscr{J} y$. Potom $y = sxt$ pro nějaké $s,t \in S^1$. Podle lemmatu 2.23 platí $sx \mathscr{L} x \mathscr{R} xt$. Díky lemmatu 2.8 z těchto ekvivalencí plyne také $sxt \mathscr{L} xt$. Dohromady dostáváme $x \mathscr{R} xt \mathscr{L} sxt$ a tedy $x \mathscr{D} y$. □

Z této věty plyne, že v konečných pologrupách můžeme mluvit o uspořádání $\mathscr{D}$-tříd, a tedy můžeme způsobem popsaným výše graficky znázorňovat strukturu nejen jednotlivých $\mathscr{D}$-tříd, ale celé pologrupy.

Příklad 2.26. Uvažujme podpologrupu S pologrupy parciálních transformací dvouprvkové množiny $\mathscr{PT}(2)$ generovanou následujícími transformacemi:

x: $\bullet \longrightarrow \bullet$ $\qquad$ y: $\bullet \longleftarrow \bullet$

Pologrupu S je rovněž možné zadat relacemi $x^2 = y^2 = 0$, $xyx = x$ a $yxy = y$. Následující obrázek popisuje Greenovy ekvivalence v této pologrupě. Přitom idempotentní prvky jsou podtrženy a je vyznačeno, jak násobení generátory přenáší mezi sebou $\mathscr{L}$-třídy a $\mathscr{R}$-třídy.

	$\xrightarrow{\cdot y}$	$\xleftarrow{\cdot x}$
$y\cdot\downarrow$	x	$\underline{xy}$
$x\cdot\uparrow$	$\underline{yx}$	y

$$\boxed{\underline{0}}$$

Příklad 2.27. Nyní uvažujme podpologrupu S pologrupy $\mathscr{PT}(2)$ generovanou transformacemi

x: $\bullet \longrightarrow \bullet$ $\qquad$ y: $\bullet \circlearrowleft \longleftarrow \bullet$

Relace zadávající tuto pologrupu jsou $x^2 = 0, xyx = x, yxy = y$ a $y^2 = y$. Obrázek popisující stejným způsobem Greenovy ekvivalence S nyní vypadá následovně:

	$\xrightarrow{\cdot y}$	$\xleftarrow{\cdot x}$
$y\cdot\downarrow$	x	$\underline{xy}$
$x\cdot\uparrow$	$\underline{yx}$	$\underline{y}$

$$\boxed{\underline{0}}$$

Všimněte si rozdílného chování násobení a odlišné polohy podgrup v horní $\mathscr{D}$-třídě oproti předchozímu příkladu.

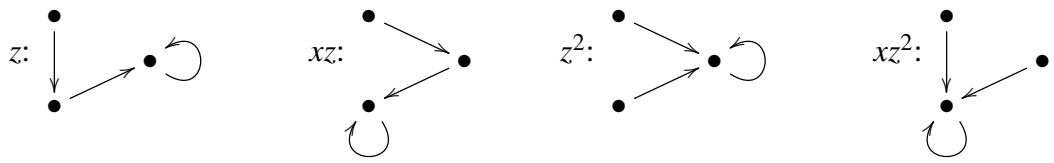
Následující cvičení ukazuje, že v nekonečných pologrupách se mohou relace $\mathscr{J}$ a $\mathscr{D}$ skutečně lišit.

Cvičení 2.28. Ukažte, že v monoidu S všech dolních trojúhelníkových matic nad kladnými racionálními čísly tvaru $\left(\begin{smallmatrix} a & 0 \\ b & 1 \end{smallmatrix}\right)$ s operací násobení je relace $\mathscr{D}$ triviální, zatímco $\mathscr{J} = S \times S$.

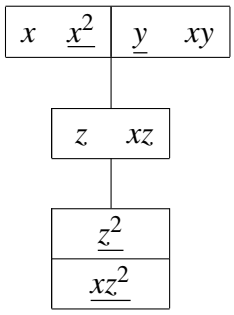
Na dalším příkladu si ukážeme, jaký je rozdíl mezi tím, když uvažujeme Greenovy ekvivalence v nějaké pologrupě, a když je uvažujeme v nějaké její podpologrupě.

Příklad 2.29. Nechť je S podpologrupa pologrupy $\mathscr{T}(3)$ sestávající z následujících prvků:

x: $\quad$ x^2: $\quad$ y: $\quad$ xy:



Greenovy ekvivalence vypadají v této pologrupě takto:



Všimněte si, že prvky x a z patří do různých $\mathscr{J}$-tříd pologrupy S, přestože patří dokonce do stejné $\mathscr{R}$-třídy pologrupy $\mathscr{T}(3)$.

Speciálním případem lemmatu 2.23 je následující důležitá vlastnost periodických pologrup, která říká, že různé $\mathscr{L}$-třídy a $\mathscr{R}$-třídy ve stejné $\mathscr{J}$-třídě jsou vždy nesrovnatelné.

Lemma 2.30. *Nechť x a y jsou prvky periodické pologrupy S, které patří do stejné $\mathscr{J}$-třídy a splňují $y \leq_{\mathscr{L}} x$. Potom platí $x \mathscr{L} y$.*

Důkaz. Protože $y \leq_{\mathscr{L}} x$, existuje $s \in S^1$ splňující $y = sx$. Použitím lemmatu 2.23 pro $t = 1$ dostáváme $y = sx \mathscr{L} x$. □

Následující dvě tvrzení jsou pouze jinými formulacemi předchozího lemmatu a uvádíme je proto, abychom si lépe uvědomili, v jakých situacích je možné toto lemma použít. Například první formulace nám říká, že v konečné pologrupě se nikdy nedokážeme dostat do jiné $\mathscr{R}$-třídy v rámci stejné $\mathscr{J}$-třídy násobením zprava.

Důsledek 2.31. *V každé periodické pologrupě z nerovnosti $x \leq_{\mathscr{J}} xs$ plyne $x \mathscr{R} xs$.*

Důsledek 2.32. *V každé periodické pologrupě z nerovnosti $x <_{\mathscr{L}} y$ plyne $x <_{\mathscr{J}} y$.*

Cvičení 2.33. V libovolném monoidu zřejmě existuje nejvyšší $\mathscr{J}$-třída vzhledem k $\leq_{\mathscr{J}}$, a to J_1. Využijte právě uvedená tvrzení k důkazu, že v periodickém monoidu je tato $\mathscr{J}$-třída tvořena právě invertibilními prvky a je podgrupou.

Z lemmatu 2.23 rovněž plyne, že pokud součin prvků konečné pologrupy zůstává ve stejné $\mathscr{J}$-třídě, potom je možné v součinu krátit:

Důsledek 2.34. *Jsou-li x, s, t a u prvky periodické pologrupy takové, že $x \mathscr{J} sxt = sxu$, potom platí $xt = xu$.*

Důkaz. Lemma 2.23 dává $sx \mathscr{L} x \mathscr{R} xt \mathscr{R} xu$. Proto podle lemmatu 2.8 definuje násobení zleva prvkem s bijekci mezi $\mathscr{R}$-třídami R_x a R_{sx}, a tedy z rovnosti $sxt = sxu$ plyne $xt = xu$. □

Dalším zajímavým důsledkem předchozích lemmat je, že pokud se obnásobením nějakého prvku konečné pologrupy současně z obou stran vrátíme do jeho $\mathscr{H}$-třídy, potom se do této $\mathscr{H}$-třídy vrátíme již po vynásobení z jedné strany.

Důsledek 2.35. *Pro libovolné prvky x, s a t periodické pologrupy splňující $sxt \in H_x$ platí $sx, xt \in H_x$.*

Důkaz. Triviálně platí $sx \leq_{\mathscr{L}} x$ a z předpokladu navíc plyne, že $sx \mathscr{J} x$ a $x \leq_{\mathscr{R}} sx$. Tedy můžeme lemma 2.30 a jeho symetrickou verzi pro relaci $\mathscr{R}$ aplikovat na prvky x a sx, čímž dostáváme požadované $sx \mathscr{L} x$ a $x \mathscr{R} sx$. V případě prvku xt můžeme postupovat analogicky. □

Lemma 2.30 můžeme také využít k důkazu, že každá podgrupa libovolného homomorfního obrazu konečné pologrupy S je homomorfním obrazem nějaké podgrupy S.

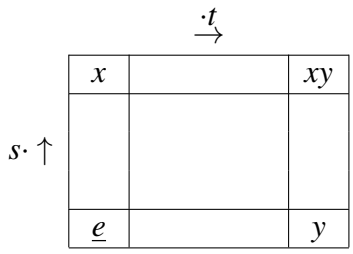
Lemma 2.36. *Je-li $\varphi\colon S \twoheadrightarrow T$ surjektivní homomorfismus konečných pologrup, potom je každá podgrupa T φ-obrazem podgrupy S. Přitom každá maximální podgrupa pologrupy T je φ-obrazem maximální podgrupy pologrupy S.*

Důkaz. Tvrzení stačí dokázat pro maximální podgrupy, což jsou podle lemmatu 2.17 právě $\mathscr{H}$-třídy obsahující idempotentní prvek. Buď tedy e libovolný idempotentní prvek pologrupy T. Nechť $x \in S$ je nějaký $\mathscr{J}$-minimální prvek v množině $\varphi^{-1}(H_e)$ a označme písmenem f idempotentní prvek x^ω. Potom $\varphi(f) = (\varphi(x))^\omega = e$. Ukážeme, že $H_e = \varphi(H_f)$. Inkluze $\varphi(H_f) \subseteq H_e$ plyne přímo z definice relace $\mathscr{H}$. Naopak, pokud $y \in S$ splňuje $\varphi(y) \in H_e$, potom $\varphi(fyf) = e \cdot \varphi(y) \cdot e = \varphi(y)$. Protože $fyf \leq_{\mathscr{J}} f \leq_{\mathscr{J}} x$, z minimality prvku x plyne, že $fyf \mathscr{J} f$. Podle důsledku 2.31 tedy platí $fyf \mathscr{H} f$. Proto $H_e \subseteq \varphi(H_f)$. □

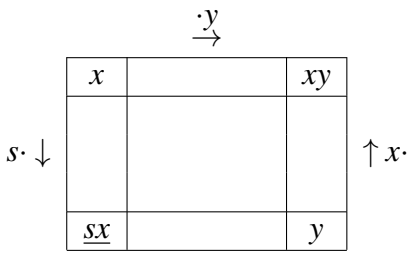
Víme, že výsledek násobení dvou prvků x a y v periodické pologrupě vždy patří do $\mathscr{D}$-třídy nižší nebo rovné D_x i D_y. Nyní se budeme zabývat otázkou, kdy je součin $\mathscr{D}$-ekvivalentních prvků x a y prvkem třídy $D_x = D_y$ a do které patří $\mathscr{H}$-třídy.

Lemma 2.37. *Nechť x a y jsou $\mathscr{J}$-ekvivalentní prvky periodické pologrupy S. Potom xy náleží do J_x právě tehdy, když existuje idempotentní prvek $e \in S$ splňující $x \mathscr{L} e \mathscr{R} y$. V tomto případě navíc platí $x \mathscr{R} xy \mathscr{L} y$.*

Důkaz. Předpokládejme nejprve, že e je idempotentní prvek splňující $x \mathscr{L} e \mathscr{R} y$. Potom existují $s,t \in S^1$ takové, že $x = se$ a $y = et$, a proto $xy = s\underline{ee}t = set = xt$, což podle lemmatu 2.8 znamená, že skutečně $x \mathscr{R} xy \mathscr{L} y$. Situaci můžeme opět znázornit obvyklým obrázkem:



Opačně, předpokládejme, že $xy \mathscr{J} x$. Protože triviálně platí $xy \leq_{\mathscr{R}} x$ a $xy \leq_{\mathscr{L}} y$, předpoklad $xy \mathscr{J} x$ podle lemmatu 2.30 znamená, že $xy \mathscr{R} x$ a současně $xy \mathscr{L} y$. Buď tedy $s \in S^1$ prvek splňující $sxy = y$. Situace v naší $\mathscr{D}$-třídě potom vypadá následovně:



Podle lemmatu 2.8 definuje násobení zleva prvky x a s vzájemně inverzní bijekce mezi $\mathscr{R}$-třídami R_y a R_x, které zachovávají $\mathscr{L}$-třídy. Proto $x \mathscr{L} sx \mathscr{R} y$ a platí $x \cdot (s \cdot x) = x$, což znamená, že prvek sx je idempotentní. □

Všimněte si, jak toto lemma souvisí s lemmatem 2.16, kde jsme se zabývali součiny dvou prvků patřících do stejné $\mathscr{H}$-třídy.

Předchozí lemma pěkně ilustruje pologrupa popsaná v příkladu 2.27: díky idempotenci prvku y musí být součin $(xy)(yx)$ roven x, zatímco součin $(yx)(xy)$ musí být roven 0, protože prvek x idempotentní není.

2.3.3 Regulární $\mathscr{D}$-třídy

V této části si ukážeme, že $\mathscr{D}$-třídy je možné rozdělit na dva základní typy: regulární, v nichž ke každému prvku existuje takzvaná pologrupová inverze, a neregulární, kde inverzi nemá žádný prvek. Regulární $\mathscr{D}$-třídy mají podstatný význam pro vlastnosti regulárních jazyků, neboť v konečných pologrupách je regulární $\mathscr{D}$-třída právě taková, že můžeme vynásobit neomezeně mnoho jejích prvků, aniž bychom tuto $\mathscr{D}$-třídu opustili, zatímco libovolný součin prvků nějaké neregulární $\mathscr{D}$-třídy vždy náleží do třídy nižší.

Říkáme, že prvek $y \in S$ je (pologrupovou) *inverzí* prvku $x \in S$, jestliže platí $xyx = x$ a $yxy = y$. Prvek pologrupy, který má inverzi, se nazývá *regulární*.

Cvičení 2.38. Dokažte, že prvek $x \in S$ je regulární právě tehdy, když existuje $y \in S$ splňující $xyx = x$.

Je-li S grupa, potom každý prvek má právě jednu pologrupovou inverzi, která se shoduje s grupovou inverzí. Proto díky lemmatu 2.17 existuje v libovolné pologrupě ke každému grupovému prvku právě jedna pologrupová inverze, která s ním leží ve stejné podgrupě; této inverzi říkáme *grupová inverze*.

Jak uvidíme dále, na rozdíl od grupových inverzí může mít jeden prvek pologrupových inverzí libovolně mnoho. Nejprve si všimněme, že jsou-li x a y vzájemně inverzní prvky, potom leží ve stejné $\mathscr{D}$-třídě, neboť $x \mathscr{R} xy \mathscr{L} y$ a $x \mathscr{L} yx \mathscr{R} y$. Přitom oba prvky xy a yx jsou idempotentní. Situace v této $\mathscr{D}$-třídě tedy vypadá následovně:

$$\begin{array}{ccc} & \xrightarrow{\cdot y} & \\ & \begin{array}{|c|c|c|} \hline x & & \underline{xy} \\ \hline & & \\ \hline \underline{yx} & & y \\ \hline \end{array} & \\ & \xleftarrow[\cdot x]{} & \end{array}$$

$y\cdot \downarrow$ $\qquad$ $\uparrow x\cdot$

Cvičení 2.39. Dokažte, že prvek $x \in S$ je regulární právě tehdy, když existují $y, z \in S$ takové, že $x = (yz)^{\omega} y$.

Příklad 2.40. Pologrupa A^{+} neobsahuje žádné regulární prvky.

Cvičení 2.41. Nechť Q je libovolná množina. Zdůvodněte, že zobrazení $\rho\colon Q \to Q$ je idempotentním prvkem pologrupy $\mathscr{T}(Q)$ právě tehdy, když pro všechna $q \in \mathrm{Im}(\rho)$ platí $\rho(q) = q$, tedy zúžení ρ na $\mathrm{Im}(\rho)$ je identita. Ukažte, že ρ je grupovým prvkem $\mathscr{T}(Q)$ právě tehdy, když $\rho|_{\mathrm{Im}(\rho)}\colon \mathrm{Im}(\rho) \to \mathrm{Im}(\rho)$ je bijekce. Dále ukažte, že každý prvek $\mathscr{T}(Q)$ je regulární.

Cvičení 2.42. Dokažte, že v monoidu všech lineárních transformací konečněrozměrného vektorového prostoru V je každý prvek regulární. S regularitou tohoto monoidu je možné se setkat v lineární algebře, kde se často využívají takzvané zobecněné inverze neboli pseudoinverze matic.

Tvrzení 2.43 (John von Neumann, 1936; D. D. Miller a Alfred Hoblitzelle Clifford, 1956)**.** *Pro libovolnou $\mathscr{D}$-třídu D v pologrupě S jsou následující podmínky ekvivalentní:*

1. *Třída D obsahuje idempotentní prvek.*

2. *Třída D obsahuje regulární prvek.*

3. *Každý prvek třídy D je regulární.*

4. *Každá $\mathscr{L}$-třída a každá $\mathscr{R}$-třída v D obsahuje idempotentní prvek.*

Důkaz. Platnost implikací $3 \implies 4 \implies 1 \implies 2$ je jasná. Zbývá tedy ukázat, že z regularity jednoho prvku dané $\mathscr{D}$-třídy plyne regularita všech ostatních. Nechť $x, y \in D$ jsou vzájemně inverzní prvky, tedy splňují $xyx = x$ a $yxy = y$. Ukážeme, že potom je regulární každý prvek $\mathscr{R}$-ekvivalentní s x. Tím bude důkaz hotov, neboť regularitu $\mathscr{L}$-ekvivalentních prvků je možné ukázat analogicky a regularita $\mathscr{D}$-ekvivalentních prvků vyplyne postupným použitím obou dokázaných faktů. Předpokládejme tedy, že z náleží do R_x. Potom existují $s, t \in S^1$ splňující $z = xs$ a $x = zt$. Ukážeme, že inverzí k z je prvek, který z vynásobením zprava převádí na idempotentní prvek $xy \in R_x$, tedy prvek ty.

		$\xrightarrow{\cdot t}$	$\xrightarrow{\cdot y}$		
	z		x		$\underline{xy = z \cdot ty}$
$y\cdot\downarrow$					
			$\underline{yx}$		y $\uparrow z\cdot$
$t\cdot\downarrow$					
	$\underline{ty \cdot z}$				ty
		$\xleftarrow[\cdot z]{}$			

Skutečně, platí

$$\underline{z \cdot ty} \cdot \underline{z} = \underline{xyxs} = xs = z,$$
$$ty \cdot \underline{z \cdot ty} = t\underline{yxy} = ty. \qquad \square$$

Definice 2.44. $\mathscr{D}$-třída se nazývá *regulární*, jestliže splňuje ekvivalentní podmínky z předchozího tvrzení.

V případě konečných pologrup dostáváme díky lemmatu 2.37 následující ekvivalentní charakterizaci regulárních $\mathscr{D}$-tříd.

Tvrzení 2.45. *Je-li S konečná pologrupa, je její $\mathscr{D}$-třída D regulární právě tehdy, když existují $x, y \in D$ takové, že $xy \in D$.*

Toto tvrzení vysvětluje význam regulárních $\mathscr{D}$-tříd v případě, kdy nás zajímá příslušnost do daného regulárního jazyka pro slova vzniklá opakovaným zřetězováním nějakých

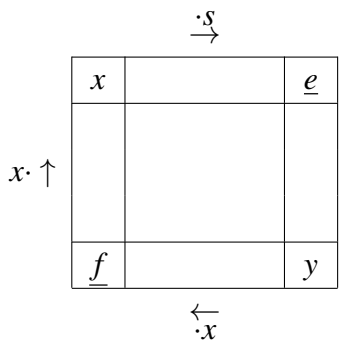
slov. Víme, že při postupném zřetězování vznikající slova reprezentují prvky ze stále nižších $\mathscr{D}$-tříd syntaktické pologrupy. Přitom pouze při zřetězování slov reprezentujících prvky regulární $\mathscr{D}$-třídy může dojít k tomu, že se budou neustále opakovat syntakticky ekvivalentní slova.

Lemma 2.46. *Každý idempotentní prvek je levý neutrální ve své $\mathscr{R}$-třídě a pravý neutrální ve své $\mathscr{L}$-třídě.*

Cvičení 2.47. Dokažte toto lemma.

Věta 2.48 (D. D. Miller a Alfred Hoblitzelle Clifford, 1956)**.** *Existuje bijekce mezi inverzemi prvku $x \in S$ a uspořádanými dvojicemi idempotentních prvků (e, f) takovými, že $e \mathscr{R} x \mathscr{L} f$. Přesněji, pro každou takovou dvojici (e, f) existuje právě jedna inverze y prvku x v $\mathscr{H}$-třídě $L_e \cap R_f$, a ta splňuje $xy = e$ a $yx = f$.*

Důkaz. Víme, že každá inverze y prvku x určuje dvojici idempotentních prvků (xy, yx) splňující $xy \mathscr{R} x \mathscr{L} yx$ a $xy \mathscr{L} y \mathscr{R} yx$. Zbývá tedy ukázat, že ke každé dvojici (e, f) taková inverze existuje a že je určená jednoznačně. Protože $e \mathscr{R} x$, platí $e = xs$ pro nějaké $s \in S^1$. Zvolme tedy za y prvek fs, který patří do $\mathscr{H}$-třídy $L_e \cap R_f$ díky lemmatu 2.8.



Potom několikanásobným použitím lemmatu 2.46 ověříme, že prvek y je skutečně inverzní k x:

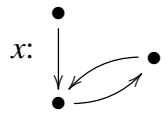
$$xyx = x\underline{fs}x = \underline{xs}x = ex = x,$$
$$yxy = y\underline{xf}s = y\underline{xs} = ye = y.$$

Předpokládejme nyní, že z je nějaká inverze prvku x v $\mathscr{H}$-třídě $L_e \cap R_f$. Protože platí $xz \mathscr{H} xy$ a $zx \mathscr{H} yx$ a každá $\mathscr{H}$-třída obsahuje podle lemmatu 2.16 nejvýše jeden idempotentní prvek, dostáváme $xz = xy$ a $zx = yx$, a tedy

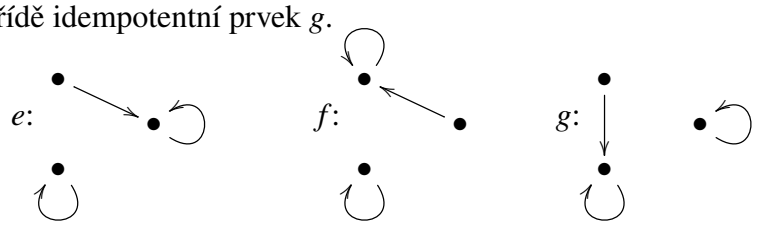
$$z = \underline{zx}z = y\underline{xz} = yxy = y.$$

Tím je dokázána jednoznačnost inverze k x ve třídě $L_e \cap R_f$. □

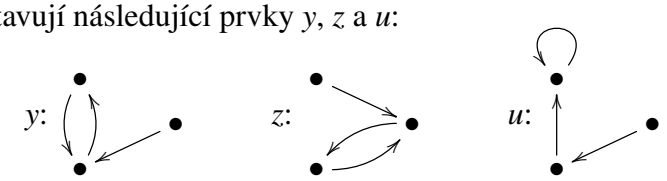
Příklad 2.49. Uvažujme $\mathscr{D}$-třídu následující transformace x v pologrupě $\mathscr{T}(3)$:



Ze cvičení 2.21 víme, že tato $\mathscr{D}$-třída obsahuje tři $\mathscr{L}$-třídy a tři $\mathscr{R}$-třídy, přičemž každá $\mathscr{H}$-třída je dvouprvková. Prvek x leží v grupové $\mathscr{H}$-třídě, jejímž idempotentním prvkem je níže uvedená transformace e. Dále leží s x ve stejné $\mathscr{L}$-třídě ještě idempotentní prvek f a ve stejné $\mathscr{R}$-třídě idempotentní prvek g.



Prvek x tedy má podle věty 2.48 právě čtyři inverze: sám sobě je grupovou inverzí a zbylé tři inverze představují následující prvky y, z a u:



Celkem tedy $\mathscr{D}$-třída prvku x vypadá následovně, přičemž čtverečky reprezentují ty její prvky, které jsme si neoznačili.

x	$\underline{e}$	z	$\underline{g}$	$\square$	$\square$
y	$\underline{f}$	u	$\square$	$\square$	$\underline{\square}$
$\square$	$\square$	$\square$	$\underline{\square}$	$\square$	$\underline{\square}$

Důsledek 2.50. *Idempotentní prvky e a f patří do stejné $\mathscr{D}$-třídy právě tehdy, když existují vzájemně inverzní prvky x a y splňující $e = xy$ a $f = yx$.*

Důkaz. Stačí zvolit libovolný prvek $x \in R_e \cap L_f$ a použít větu 2.48. $\square$

Následující lemma ukazuje, že v regulární $\mathscr{D}$-třídě je možné mezi různými prvky přecházet nejen jejich násobením nějakými prvky pologrupy S, ale dokonce přímo pomocí prvků této $\mathscr{D}$-třídy.

Lemma 2.51. *Nechť D je regulární $\mathscr{D}$-třída pologrupy S a nechť $x, y \in D$. Potom existuje prvek $z \in D$ takový, že $x \in zD$, $z \in xD$, $y \in Dz$ a $z \in Dy$.*

Důkaz. Nechť z je libovolný prvek $\mathscr{H}$-třídy $R_x \cap L_y$. Ukážeme, že platí $x \in zD$, přičemž ostatní požadované vlastnosti $z \in xD$, $y \in Dz$ a $z \in Dy$ lze dokázat analogicky. Uvažujme libovolný idempotentní prvek e v $\mathscr{R}$-třídě $R_x = R_z$, který existuje díky regularitě třídy D. Podle věty 2.48 existuje k prvku z inverze z', pro kterou platí $zz' = e$. Potom ovšem $z \cdot z'x = ex = x$ podle Lemmatu 2.46, přičemž prvek $z'x$ patří do D, neboť $z'x \mathscr{L} x$. $\square$

2.4 Hlavní faktory konečných pologrup

Nyní se budeme zabývat otázkou, jak se chová násobení lokálně v jedné $\mathscr{D}$-třídě konečné pologrupy S. Pokud tedy součin nějakých prvků této $\mathscr{D}$-třídy patří do $\mathscr{D}$-třídy nižší, bude nás pouze zajímat, že součin opustil $\mathscr{D}$-třídu, a nikoli, do které $\mathscr{D}$-třídy náleží. Takto každé $\mathscr{D}$-třídě přiřadíme nějakou pologrupu, která se nazývá hlavní faktor pologrupy S. Uvidíme, že každá $\mathscr{D}$-třída se ve skutečnosti chová jako pologrupa jednoho ze tří typů, které nyní definujeme.

Abychom mohli formálně mluvit o pologrupě odpovídající lokálnímu násobení v jedné $\mathscr{D}$-třídě, uvážíme speciální kongruence na pologrupě S, které jsou určené ztotožněním všech prvků nějakého jejího ideálu. Je-li I neprázdný ideál pologrupy S, potom je $\sim = \mathrm{id}_S \cup (I \times I)$ kongruence na S, příslušný kvocient $S/\sim$ se nazývá *Reesův kvocient* a značí se S/I. Snadno se vidí, že v pologrupě S/I všechny prvky ideálu I reprezentují nulový prvek, tedy S/I je až na izomorfismus tvaru $(S \setminus I) \cup \{0\}$, kde 0 je nový nulový prvek, přičemž součiny prvků patřících do $S \setminus I$ jsou rovny 0, pokud původně patřily do I, a jinak jsou stejné jako v S.

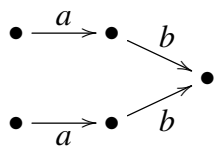
Definice 2.52. *Nulová pologrupa* (null semigroup) je pologrupa obsahující nulový prvek 0, který vznikne součinem libovolných dvou prvků této pologrupy, tedy $S^2 = \{0\}$.

Definice 2.53. Říkáme, že pologrupa je *jednoduchá*, jestliže je svým jediným neprázdným ideálem, tedy jestliže sestává z jediné $\mathscr{J}$-třídy.

Uvědomte si, že jednoduchá pologrupa nemusí být jednoduchou algebrou ve smyslu obvyklém v univerzální algebře, neboť obvyklá definice vyžaduje neexistenci netriviální kongruence, ovšem ne každá kongruence na pologrupě je definovaná ideálem.

Cvičení 2.54. Pomocí lemmatu 2.37 dokažte, že konečná pologrupa S je jednoduchá právě tehdy, když pro všechny prvky $x, y \in S$ platí $x^{\omega+1} = x$ a $(xyx)^\omega = x^\omega$. Všimněte si, že rovnost $x^{\omega+1} = x$ je ekvivalentní faktu, že prvek x patří do nějaké podgrupy S.

Cvičení 2.55. Podobně jako ve cvičení 1.19 v případě grup můžeme charakterizovat rozpoznávání jazyků konečnými jednoduchými pologrupami pomocí vlastností jejich minimálního automatu: Dokažte, že jazyk $L \subseteq A^+$ je rozpoznávaný konečnou jednoduchou pologrupou právě tehdy, když jeho minimální automat je konečný a neexistují v něm dva stavy p a q takové, že pro nějaká písmena $a, b \in A$ jsou stavy $\delta(p,a)$ a $\delta(q,a)$ různé, ale přitom platí $\delta(p,ab) = \delta(q,ab)$, tedy jestliže minimální automat neobsahuje vzor



Důležitým zobecněním jednoduchých pologrup jsou takzvané 0-jednoduché pologrupy.

Definice 2.56. Pologrupa se nazývá 0-*jednoduchá*, jestliže obsahuje nulový prvek, má právě dva neprázdné ideály (sestává tedy právě ze dvou $\mathscr{J}$-tříd $\{0\}$ a $S \setminus \{0\}$) a přitom není nulovou pologrupou (tedy $S^2 \neq \{0\}$).

Uvědomte si, že požadavek, aby 0-jednoduchá pologrupa nebyla nulová, slouží v této definici pouze k vyloučení dvouprvkových nulových pologrup, neboť větší nulové pologrupy sestávají z více než dvou $\mathscr{J}$-tříd. Příkladem 0-jednoduchých pologrup jsou například pologrupy v příkladech 2.26 a 2.27. Všimněte si, že 0-jednoduché pologrupy skutečně zobecňují jednoduché pologrupy, neboť přidáním nového nulového prvku k jednoduché pologrupě vznikne 0-jednoduchá pologrupa.

Prvním *hlavním faktorem* (principal factor) konečné pologrupy S je její nejnižší $\mathscr{D}$-třída (tedy její nejmenší neprázdný ideál), která je podle tvrzení 2.45 regulární a tedy je podle lemmatu 2.51 jednoduchou pologrupou.

Uvažujme nyní nějakou jinou $\mathscr{D}$-třídu D pologrupy S a dívejme se na ideál S^1DS^1 generovaný D jako na pologrupu. Snadno se nahlédne, že množina $S^1DS^1 \setminus D$ je neprázdným ideálem pologrupy S^1DS^1. Proto můžeme uvážit Reesův kvocient $S^1DS^1/(S^1DS^1 \setminus D)$, o kterém budeme rovněž mluvit jako o *hlavním faktoru* pologrupy S. Není-li třída D regulární, potom je podle tvrzení 2.45 jí příslušející hlavní faktor nulovou pologrupou. Je-li D regulární, potom je podle tvrzení 2.51 příslušný hlavní faktor 0-jednoduchou pologrupou.

Podívejme se nyní, jak vypadají hlavní faktory libovolné pologrupy, která je homomorfním obrazem pologrupy S.

Tvrzení 2.57. *Je-li $\varphi\colon S \twoheadrightarrow T$ surjektivní homomorfismus konečných pologrup, potom je každý hlavní faktor pologrupy T homomorfním obrazem nějakého hlavního faktoru pologrupy S.*

Důkaz. Nechť D je libovolná $\mathscr{D}$-třída pologrupy T. Mezi prvky množiny $\varphi^{-1}(D)$ zvolme libovolný $\mathscr{J}$-minimální prvek $x \in S$. Potom platí, že $\mathscr{J}$-třída D_x je nejnižší v S právě tehdy, když $\mathscr{J}$-třída D je nejnižší v T. K tomu, abychom ukázali, že hlavní faktor příslušný třídě D je obrazem hlavního faktoru příslušného třídě D_x v homomorfismu indukovaném φ, tedy stačí ověřit, že D je φ-obrazem třídy D_x a že φ všechny prvky ideálu $S^1D_xS^1 \setminus D_x$ zobrazí do ideálu $T^1DT^1 \setminus D$.

Protože libovolné $\mathscr{D}$-ekvivalentní prvky se vždy zobrazí na $\mathscr{D}$-ekvivalentní prvky, platí $\varphi(D_x) \subseteq D$. Pro ověření opačné inkluze vezměme libovolný prvek $t \in D$. Protože $t \mathrel{\mathscr{J}} \varphi(x)$, existují $s, u \in S^1$ splňující $t = \varphi(s)\varphi(x)\varphi(u) = \varphi(sxu)$. Zřejmě platí $sxu \leq_{\mathscr{J}} x$. Ovšem $sxu <_{\mathscr{J}} x$ platit nemůže díky $\mathscr{J}$-minimalitě prvku x. Proto sxu patří do D_x a tedy $t \in \varphi(D_x)$.

Pokud y je prvek ideálu $S^1D_xS^1$, který nepatří do D_x, potom $y <_{\mathscr{J}} x$ a proto $\varphi(y) \leq_{\mathscr{J}} \varphi(x)$, přičemž $\varphi(y) \mathrel{\mathscr{J}} \varphi(x)$ nemůže nastat díky volbě x. $\square$

2.4.1 Konečné 0-jednoduché pologrupy

Naším cílem nyní bude popsat všechny konečné 0-jednoduché pologrupy. Nejprve si všimněme dvou speciálních případů jednoduchých pologrup: Každá grupa je jednoduchou pologrupou s jedinou $\mathscr{H}$-třídou, a na druhou stranu, takzvané rektangulární bandy, které si nyní definujeme, jsou jednoduché pologrupy, které mají všechny $\mathscr{H}$-třídy triviální.

Definice 2.58. Nechť R a L jsou libovolné množiny. Potom pologrupa definovaná na množině $R \times L$ operací $(r,\ell) \cdot (r',\ell') = (r,\ell')$ se nazývá *rektangulární band*.

Poznámka 2.59. V teorii pologrup se termínem *band* označuje libovolná idempotentní pologrupa.

Všimněte si, že v rektangulárním bandu patří prvky (r,ℓ) a (r',ℓ') do stejné $\mathscr{R}$-třídy právě tehdy, když $r = r'$ a do stejné $\mathscr{L}$-třídy právě tehdy, když $\ell = \ell'$.

Cvičení 2.60. Dokažte, že pologrupa S je rektangulární band právě tehdy, když všechny její prvky jsou k sobě vzájemně inverzní, tedy pro všechna $x, y \in S$ platí $xyx = x$.

Následující cvičení ukazuje, že každá konečná jednoduchá pologrupa je jistým způsobem složena z rektangulárního bandu a grupy.

Cvičení 2.61. Použitím lemmatu 2.37 dokažte, že pro každou konečnou jednoduchou pologrupu S je relace $\mathscr{H}$ kongruence, kvocient $S/\mathscr{H}$ je rektangulární band a všechny $\mathscr{H}$-třídy jsou vzájemně izomorfní grupy.

Nyní si popíšeme způsob, jak lze libovolnou 0-jednoduchou pologrupu z rektangulárního bandu a grupy vytvořit. Nechť R a L jsou libovolné konečné množiny, G libovolná konečná grupa a $0 \notin G$ nějaký nový prvek. Nechť dále $P = (p_{\ell r})_{\ell \in L,\ r \in R}$ je nějaká $L \times R$-matice, jejíž prvky patří do množiny $G \cup \{0\}$, a která obsahuje v každém řádku a v každém sloupci alespoň jeden nenulový prvek. Na množině $(R \times G \times L) \cup \{0\}$ zavedeme násobení předpisem

$$(r,g,\ell) \cdot (r',g',\ell') = \begin{cases} (r, g \cdot p_{\ell r'} \cdot g', \ell'), & \text{pokud } p_{\ell r'} \neq 0, \\ 0, & \text{pokud } p_{\ell r'} = 0, \end{cases}$$

přičemž 0 je nulový prvek. Toto násobení je asociativní a vzniklá pologrupa se značí $\mathfrak{M}^0(R,L,G,P)$ a nazývá se *Reesova maticová pologrupa*. Skutečně se totiž jedná o pologrupu matic, neboť prvek (r,g,ℓ) pologrupy $\mathfrak{M}^0(R,L,G,P)$ můžeme chápat jako matici s jediným nenulovým prvkem g na pozici (r,ℓ), přičemž použijeme takzvané sendvičové násobení matic: $M \cdot N = MPN$.

Věta 2.62 (Anton Suschkewitsch, 1928; David Rees, 1940)**.** *Konečná pologrupa je 0-jednoduchá právě tehdy, když je izomorfní pologrupě $\mathfrak{M}^0(R,L,G,P)$ pro nějaké konečné množiny R a L, konečnou grupu G a matici P.*

Důkaz. Snadno se ověří, že pologrupa $\mathfrak{M}^0(R,L,G,P)$ je vždy 0-jednoduchá. Na druhou stranu, je-li S konečná 0-jednoduchá pologrupa, je díky předpokladu $S^2 \neq \{0\}$ a tvrzení 2.45 její nenulová $\mathscr{D}$-třída regulární. Za G zvolíme libovolnou grupovou $\mathscr{H}$-třídu v této $\mathscr{D}$-třídě. Označme e neutrální prvek G. Za L (respektive R) vezmeme množinu všech $\mathscr{L}$-tříd (respektive $\mathscr{R}$-tříd) v nenulové $\mathscr{D}$-třídě. Dále zvolíme pro každé $\ell \in L$ libovolný prvek $s_\ell \in \ell \cap R_e$ v příslušné $\mathscr{L}$-třídě, který leží ve stejné $\mathscr{R}$-třídě jako G. Analogicky zvolíme pro každé $r \in R$ prvek $t_r \in r \cap L_e$ v příslušné $\mathscr{R}$-třídě. Lemma 2.46 říká, že platí $e \cdot s_\ell = s_\ell$ a $t_r \cdot e = t_r$, takže podle lemmatu 2.8 násobení prvky s_ℓ a t_r zadává bijekce mezi $\mathscr{H}$-třídami G, $\ell \cap R_e$, $r \cap L_e$ a $\ell \cap r$.

		$\xrightarrow{\cdot s_\ell}$	ℓ
	G		s_ℓ
$t_r \cdot \downarrow$			
r	t_r		$t_r g s_\ell$

Jelikož každý nenulový prvek pologrupy S leží v nějaké $\mathscr{H}$-třídě $\ell \cap r$, je možné všechny nenulové prvky S jednoznačně vyjádřit ve tvaru $t_r g s_\ell$, kde $r \in R$, $g \in G$ a $\ell \in L$. Protože součinem dvou prvků tohoto tvaru je prvek $(t_r g s_\ell)(t_{r'} g' s_{\ell'}) = t_r(g s_\ell t_{r'} g') s_{\ell'}$, stačí nyní volit za prvky matice P prvky $p_{\ell r} = s_\ell t_r$, přičemž tyto prvky skutečně buď patří do G nebo jsou rovny 0 podle lemmatu 2.37. Poněvadž každá $\mathscr{L}$-třída a každá $\mathscr{R}$-třída regulární $\mathscr{D}$-třídy obsahuje nějaký idempotentní prvek, zaručuje toto lemma rovněž, že v každém řádku a v každém sloupci matice P bude alespoň jeden nenulový prvek, jak se v konstrukci pologrupy $\mathfrak{M}^0(R,L,G,P)$ požaduje. □

Všimněte si, že předchozí věta dává rovněž charakterizaci konečných jednoduchých pologrup; tyto pologrupy vzniknou odebráním prvku 0 z těch pologrup $\mathfrak{M}^0(R,L,G,P)$, kde všechny prvky matice P patří do G.

Poznámka 2.63. Důležitou konstrukcí, kterou můžeme zavést podobně jako výše popsanou reprezentaci 0-jednoduchých pologrup, je *Schützenbergerho reprezentace* libovolné pologrupy S odpovídající zvolené $\mathscr{D}$-třídě D této pologrupy. Každý prvek $s \in S$ se reprezentuje jako čtvercová matice, jejíž řádky i sloupce odpovídají $\mathscr{L}$-třídám v D, přičemž každý řádek obsahuje nejvýše jeden nenulový prvek, který vyjadřuje, na kterou $\mathscr{L}$-třídu L' prvek s převádí $\mathscr{L}$-třídu L, jíž tento řádek odpovídá. Onen nenulový prvek je prvkem Schützenbergerho grupy nějaké pevně zvolené $\mathscr{L}$-třídy $\tilde{L}$ v D a získáme jej tak, že složíme bijekci mezi L a L' danou prvkem s s pevně zvolenými bijekcemi mezi třídou $\tilde{L}$ a třídami L a L'.

Schützenbergerho reprezentace tedy kóduje informaci, o kterou přicházíme, pokud studujeme každou $\mathscr{D}$-třídu samostatně jako hlavní faktor: říká nám, jak na naší $\mathscr{D}$-třídě působí prvky z vyšších $\mathscr{D}$-tříd.

2.5 Příklady aplikací

Nyní si ukážeme dva výsledky, které demonstrují úzkou vazbu mezi vlastnostmi regulárních jazyků a strukturou Greenových relací jejich syntaktických pologrup. Většina konstrukcí a důkazů je u výsledků tohoto typu založena na vypořádání se se situací v jedné $\mathscr{D}$-třídě a použití indukce vzhledem k uspořádání $\mathscr{D}$-tříd shora dolů, která odpovídá postupnému prodlužování slov.

2.5.1 Star-free jazyky a aperiodické pologrupy

Definice 2.64. Regulární jazyk $L \subseteq A^*$ se nazývá *star-free*, jestliže jej lze definovat pomocí racionálního výrazu používajícího místo Kleeneho iterace operaci komplementu v A^*.

Povolenými symboly při tvorbě star-free výrazů jsou tedy písmena $a \in A$, $\emptyset$, sjednocení, zřetězení a komplement $\bar{\cdot}$. Všimněte si, že komplement přebírá od iterace úlohu jediné operace schopné vytvořit z konečných jazyků jazyky nekonečné. Díky De Morganovým zákonům můžeme samozřejmě ve star-free výrazech používat i operaci průniku.

Cvičení 2.65. Nechť A je abeceda a $B \subseteq A$ její podabeceda. Ukažte, že jazyk B^* je star-free nad abecedou A. Použijte tohoto faktu k důkazu, že star-free jazyky nad abecedou B jsou právě star-free jazyky nad abecedou A obsažené v B^*. Vlastnost jazyka „být star-free“ tedy nezávisí na volbě abecedy, nad kterou tento jazyk uvažujeme.

Jak ukazuje následující příklad, star-free výraz je možné zkonstruovat i pro jazyky, u kterých není na první pohled zřejmé, jak se použití iterace vyhnout.

Příklad 2.66. Jazyk $(ab)^+$ nad abecedou $\{a,b\}$ je star-free, neboť lze zadat výrazem

$$a\bar{\emptyset} \cap \bar{\emptyset}b \cap \overline{\bar{\emptyset}aa\bar{\emptyset}} \cap \overline{\bar{\emptyset}bb\bar{\emptyset}}.$$

Tento výraz totiž říká, že slovo patří do jazyka $(ab)^+$ právě tehdy, když začíná na a, končí na b a písmena a a b se v něm pravidelně střídají.

Není tedy snadné algoritmicky určit, zda daný regulární jazyk je možné zadat star-free výrazem. Tento problém se podařilo vyřešit dokázáním, že star-free jazyky jsou právě jazyky rozpoznatelné konečnými aperiodickými monoidy.

Definice 2.67. Pologrupa se nazývá *aperiodická*, jestliže je periodická a všechny její podgrupy jsou triviální, tedy jednoprvkové.

Třídu aperiodických pologrup můžeme mezi konečnými pologrupami chápat v podstatě jako „ortogonální doplněk“ ke třídě všech konečných grup.

Cvičení 2.68. Dokažte, že periodická pologrupa S je aperiodická právě tehdy, když pro všechna $x \in S$ platí $x^{\omega+1} = x^{\omega}$.

Uvědomte si, že pologrupa S transformací konečné množiny Q je aperiodická právě tehdy, když pro každou transformaci $f \in S$ existuje přirozené číslo n takové, že pro všechna $q \in Q$ platí $f^n(q) = f^{n+1}(q)$.

Cvičení 2.69. Dokažte, že jazyk $L \subseteq A^*$ je rozpoznávaný konečnou aperiodickou pologrupou právě tehdy, když jeho minimální automat je konečný a neexistuje v něm cyklus ohodnocený nějakým slovem w^n, kde $w \in A^+$ a $n \geq 2$ je nějaké přirozené číslo.

Aperiodicita pologrup je ve skutečnosti ekvivalentní zdánlivě silnější podmínce $\mathscr{H}$-triviality.

Tvrzení 2.70. *Periodická pologrupa S je aperiodická právě tehdy, když $\mathscr{H} = \mathrm{id}_S$.*

Důkaz. Je-li relace $\mathscr{H}$ triviální, potom jsou podle lemmatu 2.17 všechny podgrupy triviální, a tedy se jedná o aperiodickou pologrupu.

Naopak, předpokládejme, že x a y jsou dva $\mathscr{H}$-ekvivalentní prvky aperiodické pologrupy S. Potom jistě existují prvky $s,t \in S^1$ splňující $y = xs$ a $x = ty$. Proto můžeme použitím obvyklého přechodu k idempotentním prvkům spočítat

$$y = tys = t^{\omega}ys^{\omega} = t^{\omega+1}ys^{\omega} = ty = x. \qquad \square$$

Věta 2.71 (Marcel-Paul Schützenberger, 1965)**.** *Regulární jazyk $L \subseteq A^*$ je star-free právě tehdy, když jeho syntaktický monoid M_L je aperiodický.*

Jelikož podmonoidy a homomorfní obrazy aperiodických monoidů jsou triviálně aperiodické, stačí k ověření, že jazyk je star-free, dokázat, že je rozpoznávaný nějakým konečným aperiodickým monoidem.

Teprve s pomocí věty 2.71 můžeme dát příklad regulárního jazyka, který není star-free.

Příklad 2.72. Jazyk $(aa)^*$ není star-free, neboť jeho syntaktickým monoidem je dvouprvková grupa. Všimněte si rozdílu mezi tímto jazykem a jazykem z příkladu 2.66.

Schützenbergerho věta tedy mimo jiné říká, že ne každý regulární jazyk je star-free, což znamená, že komplement nedokáže plně nahradit Kleeneho iteraci. Můžeme se ovšem dále ptát, kolik v sobě zanořených iterací v rozšířeném regulárním výrazu, který používá kromě iterace i operaci komplementu, k definování libovolného regulárního jazyka potřebujeme. Dá se ukázat, že pokud používáme obyčejné racionální výrazy bez komplementu, je neomezené zanořování iterací nezbytné. Na druhou stranu, při použití komplementu dosud není ani známo, zda je vůbec nutné použít dvě v sobě zanořené iterace.

Cvičení 2.73. Dokažte přímou implikaci věty 2.71.

Abychom dokázali opačnou implikaci Schützenbergerho věty, vyjádříme v následujících lemmatech, které součiny prvků syntaktického monoidu patří do dané $\mathscr{H}$-třídy, a to pomocí povolených racionálních operací, které aplikujeme na výrazy charakterizující rovnost prvkům ve vyšších $\mathscr{J}$-třídách.

Lemma 2.74. *Pro libovolné $n \in \mathbb{N}$ a libovolné prvky $x, x_1, \ldots, x_n$ konečného monoidu S, kde x nepatří do nejvyšší $\mathscr{J}$-třídy J_1, platí $x_1 \cdots x_n \in H_x$ právě tehdy, když jsou splněny následující tři podmínky:*

1. *pro nějaké $i \in \{1, \ldots, n\}$ platí $x_1 \cdots x_i \in R_x$;*
2. *pro nějaké $i \in \{1, \ldots, n\}$ platí $x_i \cdots x_n \in L_x$;*
3. *platí $x_1 \cdots x_n \geq_{\mathscr{J}} x$.*

Cvičení 2.75. Dokažte toto lemma. Všimněte si, že předpoklad $x \notin J_1$ nebyl v důkazu použit; budeme jej totiž potřebovat až při přeformulovávání podmínek 1 až 3 v následujících lemmatech.

Tyto tři podmínky nyní vyjádříme pomocí charakterizací součinů reprezentujících prvky vyšších $\mathscr{J}$-tříd tak, že budeme uvažovat ty indexy i, pro které jsou poprvé splněny (v případě podmínek 1 a 2), respektive nesplněny (v případě podmínky 3). Přitom si vystačíme s povolenými racionálními operacemi, neboť budeme muset popsat pouze jeden krátký úsek v součinu $x_1 \cdots x_n$ (pro podmínky 1 a 2), případně dva úseky (pro podmínku 3).

Všimněte si, že ve formulacích následujících lemmat se vyskytují součiny nulové délky, které vždy chápeme jako neutrální prvek monoidu S.

Lemma 2.76. *Podmínka 1 lemmatu 2.74 je splněna právě tehdy, když pro nějaké $i \in \{1, \ldots, n\}$ platí $x_1 \cdots x_{i-1} >_{\mathscr{J}} x$ a současně $x_1 \cdots x_i \in R_x$.*

Podmínka 2 lemmatu 2.74 je splněna právě tehdy, když pro nějaké $i \in \{1, \ldots, n\}$ platí $x_{i+1} \cdots x_n >_{\mathscr{J}} x$ a současně $x_i \cdots x_n \in L_x$.

Cvičení 2.77. Dokažte toto lemma.

Lemma 2.78. *Podmínka 3 lemmatu 2.74 není splněna právě tehdy, když je splněna aspoň jedna z následujících podmínek:*

1. *pro nějaké $i \in \{1, \ldots, n\}$ platí $x_i \not\geq_{\mathscr{J}} x$;*
2. *pro nějaká $i, j \in \{1, \ldots, n\}$, $i < j$, platí $x_i \cdots x_j \not\geq_{\mathscr{J}} x$ a současně $x_{i+1} \cdots x_{j-1} >_{\mathscr{J}} x$.*

Důkaz. Je jasné, že z platnosti kterékoli ze dvou podmínek plyne, že podmínka 3 lemmatu 2.74 není splněna.

Abychom ukázali opačnou implikaci, předpokládejme, že $x_1 \cdots x_n \not\geq_{\mathscr{J}} x$ a zvolme $i, j \in \{1, \ldots, n\}$, $i \leq j$, takové, že $x_i \cdots x_j \not\geq_{\mathscr{J}} x$ a přitom rozdíl $j - i$ je nejmenší možný. Pokud $i = j$, tak je splněna první podmínka. Pokud $i < j$, označme $y = x_{i+1} \cdots x_{j-1}$. K ověření druhé podmínky stačí dokázat $y >_{\mathscr{J}} x$, čehož dosáhneme sporem. Jestliže totiž $y >_{\mathscr{J}} x$ neplatí, patří díky volbě indexů i a j prvky y, $x_i y$ a $y x_j$ do stejné $\mathscr{J}$-třídy jako x. Proto podle důsledku 2.31 platí $x_i y \mathscr{L} y$ a $y x_j \mathscr{R} y$. Můžeme tedy aplikovat lemma 2.8, díky němuž zadává násobení prvkem x_i zleva bijekci mezi R_y a $R_{x_i y}$. Situace v $\mathscr{D}$-třídě D_x je tedy následující:

$$\begin{array}{cccc} & & \xrightarrow{\cdot x_j} & \\ & \begin{array}{|c|} \hline y \\ \hline \\ \hline x_i y \\ \hline \end{array} & \begin{array}{|c|} \hline \\ \hline \\ \hline \\ \hline \end{array} & \begin{array}{|c|} \hline y x_j \\ \hline \\ \hline x_i y x_j \\ \hline \end{array} \\ x_i \cdot \downarrow & & & \end{array}$$

Proto $x_i y x_j \mathscr{J} y \mathscr{J} x$, což je ovšem ve sporu s předpokladem $x_i y x_j \not\geq_{\mathscr{J}} x$. □

Důkaz opačné implikace věty 2.71. Nechť $\varphi \colon A^* \to S$ je homomorfismus do konečného aperiodického monoidu. Stačí nám ukázat, že pro každý prvek $x \in S$ je jazyk $\varphi^{-1}(x)$ star-free. Budeme postupovat indukcí shora dolů vzhledem k předuspořádání $\geq_{\mathscr{J}}$.

Nejvyšší $\mathscr{J}$-třída monoidu S podle cvičení 2.33 obsahuje pouze neutrální prvek 1. Na tento prvek se tedy zobrazí právě slova, jejichž všechna písmena se zobrazují na 1. Proto

$$\varphi^{-1}(1) = A^* \setminus (A^* \cdot \{a \in A \mid \varphi(a) \neq 1\} \cdot A^*)$$

je star-free jazyk.

Nechť nyní $x \in S$ nepatří do nejvyšší $\mathscr{J}$-třídy S. Protože podle tvrzení 2.70 je S $\mathscr{H}$-triviální, platí $\varphi^{-1}(x) = \varphi^{-1}(H_x)$. Proto můžeme jazyk $\varphi^{-1}(x)$ vyjádřit pomocí lemmat 2.74, 2.76 a 2.78 následovně:

$$\varphi^{-1}(x) = (RA^* \cap A^* L) \setminus A^* J A^*,$$

kde

$$\begin{aligned} R &= \bigcup \{\varphi^{-1}(y) a \mid y \in S,\ a \in A,\ y >_{\mathscr{J}} x,\ y\varphi(a) \in R_x\} \\ L &= \bigcup \{a \varphi^{-1}(y) \mid y \in S,\ a \in A,\ y >_{\mathscr{J}} x,\ \varphi(a) y \in L_x\} \\ J &= \{a \in A \mid \varphi(a) \not\geq_{\mathscr{J}} x\} \\ &\quad \cup \bigcup \{a \varphi^{-1}(y) b \mid y \in S,\ a, b \in A,\ y >_{\mathscr{J}} x,\ \varphi(a) y \varphi(b) \not\geq_{\mathscr{J}} x\} \end{aligned}$$

Jelikož všechny jazyky $\varphi^{-1}(y)$, kde $y >_{\mathscr{J}} x$, jsou star-free podle indukčního předpokladu, vidíme, že je star-free i jazyk $\varphi^{-1}(x)$. □

2.5.2 Opakování prvků v součinech a vlastnost konečné mocniny*

Říkáme, že jazyk $L \subseteq A^*$ má *vlastnost konečné mocniny*, jestliže pro nějaké $n \in \mathbb{N}$ platí $L^+ = L \cup L^2 \cup \cdots \cup L^n$, tedy jestliže se iterace tohoto jazyka po konečném počtu kroků zastaví.

Věta 2.79 (Kosaburo Hashiguchi 1979; Imre Simon 1978)**.** *Je možné algoritmicky rozhodovat, zda daný regulární jazyk má vlastnost konečné mocniny.*

Algoritmus Imreho Simona je založen na rozhodování omezenosti automatů, jejichž hrany mají přiřazeny váhy v polookruhu $(\mathbb{N} \cup \{0, \infty\}, \min, +)$, zatímco Hashiguchi použil k důkazu komplikovaný a neprůhledný kombinatorický argument založený na Dirichletově zásuvkovém principu. Nyní si ukážeme, jaká je algebraická podstata Hashiguchiho argumentu. Základním nástrojem je následující lemma, které ukazuje, že pokud vynásobíme dostatečně mnoho prvků téže $\mathscr{J}$-třídy, aniž bychom tuto $\mathscr{J}$-třídu opustili, potom se některý z použitých prvků objeví jako výsledek nějakého delšího faktoru tohoto součinu.

Lemma 2.80. *Nechť J je libovolná $\mathscr{J}$-třída konečné pologrupy S a $x_1, \ldots, x_n \in S$. Jestliže $x_1 \cdots x_n \in J$ a množina indexů $\{i \in \{1, \ldots, n\} \mid x_i \in J\}$ obsahuje více prvků než třída J, potom existují $i, j \in \{1, \ldots, n\}$, $i < j$, takové, že $x_i, x_j \in J$ a $x_i \cdots x_j = x_i$.*

Důkaz. Buď $k \in \{1, \ldots, n\}$ nejmenší index takový, že $x_k \in J$. Díky předpokladu $x_1 \cdots x_n \in J$ víme, že pro všechna $j \in \{k, \ldots, n\}$ platí $x_k \cdots x_j \in J$. Jelikož $|\{i \in \{1, \ldots, n\} \mid x_i \in J\}| > |J|$, musí podle Dirichletova zásuvkového principu existovat $i, j \in \{k, \ldots, n\}$, $i < j$, která splňují $x_i, x_j \in J$ a

$$(x_k \cdots x_{i-1}) \cdot x_i = (x_k \cdots x_{i-1}) \cdot x_i \cdot (x_{i+1} \cdots x_j).$$

Důsledek 2.34 nám ovšem umožňuje v tomto součinu krátit, a tedy platí $x_i = x_i \cdots x_j$. □

Jelikož nás zajímá příslušnost slov jak do jazyka L, tak do jeho iterace, budeme uvažovat homomorfismus $\varphi \colon A^* \to S$ rozpoznávající jazyky L, L^+ a $\{\varepsilon\}$. Protože se musíme zabývat rozklady každého slova na součin slov z L, nestačí tentokrát použít přímo tento homomorfismus, ale musíme v homomorfismu podchytit rovněž rozklady slov na faktory. Toto nám zajistí následující užitečná konstrukce, kterou pravděpodobně poprvé použili Jean-Camille Birget a John L. Rhodes v roce 1984. Definujme zobrazení $\tau \colon A^* \to \mathscr{P}(S \times S \times S)$ předpisem

$$\tau(w) = \{(\varphi(t), \varphi(u), \varphi(v)) \mid t, u, v \in A^*,\ w = tuv\}.$$

Každé slovo tedy rozložíme všemi možnými způsoby na tři faktory a přiřadíme mu množinu všech takto získaných trojic prvků pologrupy S. Definujme relaci ekvivalence $\sim$ na A^* jako jádro zobrazení τ, tedy $u \sim v \iff \tau(u) = \tau(v)$.

Cvičení 2.81. Dokažte, že $\sim$ je kongruencí monoidu A^*.

Nyní tedy víme, že $\tau(A^*) \cong A^*/\sim$ je pologrupa a $\tau\colon A^* \to \tau(A^*)$ je homomorfismus. Uvědomte si ale, že námi zavedená operace na $\tau(A^*)$ je zcela odlišná od obvyklé operace na množině $\mathscr{P}(S \times S \times S)$, jejíž je $\tau(A^*)$ podmnožinou. Protože nás při rozkládání slov zajímají pouze slova patřící do jazyka L^+, posledním krokem konstrukce bude omezení pologrupy $\tau(A^*)$ na její podpologrupu $\tau(L^+)$, kterou označíme T.

Indukcí vzhledem k uspořádání $\mathscr{J}$-tříd této pologrupy T je nyní možné ukázat následující tvrzení, které charakterizuje vlastnost konečné mocniny podle toho, jak často se v pologrupě T vyskytují prvky reprezentované slovy z jazyka L. Všimněte si, že rozhodující roli hrají regulární $\mathscr{D}$-třídy, neboť pro dosažení vlastnosti konečné mocniny musíme být schopni nahrazovat dlouhé součiny mnoha slov z L, které zůstávají stále ve stejné $\mathscr{D}$-třídě, jediným slovem patřícím do L.

Tvrzení 2.82 (Michal Kunc, 2006)**.** *Pro regulární jazyk L jsou následující podmínky ekvivalentní:*

1. *Jazyk L má vlastnost konečné mocniny.*
2. *Každá regulární $\mathscr{D}$-třída pologrupy T obsahuje nějaký prvek množiny $\tau(L)$.*
3. *Platí $L^+ = L \cup \cdots \cup L^{(j+1)^h}$, kde j značí maximální velikost $\mathscr{J}$-třídy v pologrupě S a h značí délku nejdelšího řetězce $\mathscr{J}$-tříd v pologrupě T.*

2.6 Faktorizační lesy

V důkazu věty 2.71 jsme viděli, jak lze vlastnosti Greenových relací využít ke konstrukci racionálních výrazů speciálního tvaru, které popisují jazyk rozpoznávaný danou pologrupou. Jiným nástrojem, který umožňuje tvořit určité speciální racionální výrazy ze syntaktického homomorfismu, jsou faktorizační lesy. Tato technika je založena na faktu, že v hodně dlouhých součinech prvků konečné pologrupy je vždy mnoho úseků, jejichž součinem je idempotentní prvek.

Nejprve si všimněme, že výsledek každého součinu, jehož délka je aspoň taková jako počet prvků pologrupy, lze zapsat použitím nějakého idempotentního prvku.

Lemma 2.83. *Je-li S konečná pologrupa, potom pro všechna $n \geq |S|$ platí $S^n = S \cdot E(S) \cdot S$.*

Důkaz. Inkluze $S^n \supseteq S \cdot E(S) \cdot S$ se pro $n \leq 2$ snadno ověří a pro $n \geq 3$ plyne z toho, že je-li $e \in E(S)$, pak $xey = xe^{n-2}y$.

Za účelem ověření opačné inkluze uvažme libovolné prvky $x_1, \ldots, x_n \in S$. Jsou-li všechny součiny $x_1 \cdots x_i$ pro $i = 1, \ldots, n$ různé, potom některý z nich musí být idempotentní a součin $x_1 \cdots x_n$ můžeme psát jako $(x_1 \cdots x_i)^2 (x_1 \cdots x_n) \in S \cdot E(S) \cdot S$. Pokud se

některé z těchto součinů rovnají, tedy existují $1 \leq i < j \leq n$ splňující $x_1 \cdots x_i = x_1 \cdots x_j$, potom

$$x_1 \cdots x_n = (x_1 \cdots x_i) \cdot (x_{i+1} \cdots x_j)^{\omega} \cdot (x_{i+1} \cdots x_n) \in S \cdot E(S) \cdot S. \qquad \square$$

Pokud navíc požadujeme, aby použitý idempotentní prvek vznikl přímo vyhodnocením nějakého faktoru našeho součinu, dá se dokázat, že stačí požadovat délku součinu alespoň 2^n, kde n je počet neidempotentních prvků v pologrupě S, přičemž tato hodnota je nejnižší možná (Thomas Eric Hall a Mark V. Sapir, 1996).

Vytvořme nyní pro libovolný součin w prvků pologrupy S binární strom s kořenem ohodnoceným w, jehož všechny uzly jsou ohodnoceny faktory w, přičemž ohodnocení přímých následníků každého uzlu ohodnoceného nějakým součinem v tvoří rozklad v přibližně na poloviny. Všimněte si, že má-li součin w délku nejvýše $2^{|S|}$, je délka každé větve takto vzniklého stromu omezená počtem prvků pologrupy S. Cílem zavedení faktorizačních lesů je získat rovněž pro delší součiny stromy rozkladů, jejichž všechny větve mají délku omezenou konstantou, kterou lze spočítat z velikosti pologrupy S. To nám potom umožní například konstruovat racionální výrazy pro jazyky indukcí vzhledem k výšce stromu, podobně jako jsme použili indukci vzhledem k uspořádání $\mathscr{J}$-tříd v důkazu věty 2.71, protože po konečně mnoha aplikacích indukčního kroku již budou do konstrukce zahrnuta všechna slova libovolného jazyka. Přitom musíme samozřejmě nejen upustit od požadavku, aby strom byl binární, ale musíme dokonce povolit větvení na neomezený počet následníků. Takové větvení ovšem povolíme pouze ve velmi speciálních případech, abychom zachovali užitečné vlastnosti vzniklého stromu. Tímto případem bude opakovaný výskyt téhož idempotentního prvku e těsně za sebou: protože víme, že součin libovolně mnoha kopií e dává vždy výsledek e, neztrácíme použitím takového rozkladu žádnou informaci o součinech některých faktorů. Přitom skutečně platí, že ve velmi dlouhých součinech se vždy nějaký idempotentní prvek musí mnohokrát těsně po sobě zopakovat:

Tvrzení 2.84. *Pro každou konečnou pologrupu S a libovolné $k \in \mathbb{N}$, $k \geq 2$, existuje $n \in \mathbb{N}$ takové, že pro libovolné prvky $x_1, \ldots, x_n \in S$ existuje idempotentní prvek $e \in E(S)$ a indexy $i_1, \ldots, i_k \in \mathbb{N}_0$, $0 \leq i_1 < \cdots < i_k \leq n$, splňující $x_{i_\ell+1} \cdots x_{i_m} = e$ pro všechna $\ell, m \in \{1, \ldots, k\}$ taková, že $\ell < m$.*

Důkaz. Chápejme S jako množinu barev. Uvažme úplný graf G na množině uzlů $\{0, \ldots, n\}$ a označme každou hranu $\{i, j\}$ barvou $x_{i+1} \cdots x_j$. Je-li $k \geq 3$, potom je existence idempotentního prvku s požadovanými vlastnostmi ekvivalentní existenci úplného podgrafu grafu G o k uzlech, jehož všechny hrany jsou ohodnocené stejnou barvou. Proto existence hledaného čísla n vyplývá přímo z Ramseyho věty. $\square$

Všimněte si, že předchozí důkaz ve skutečnosti vůbec nevyužívá asociativity násobení v pologrupě. V dalším textu ukážeme, že s využitím asociativity je možné tvrzení 2.84 výrazně zesílit: výskyty mnoha kopií téhož idempotentního prvku těsně za sebou jsou

totiž natolik časté, že umožňují výše popsaným způsobem rozložit libovolně dlouhý součin prvků pologrupy S na jednotlivé prvky pomocí stromu, jehož výška je omezená konstantou závisející pouze na pologrupě S. Formálně přitom budeme místo se součiny prvků S pracovat se součiny písmen, která budou vyhodnocována v pologrupě S pomocí nějakého pevného homomorfismu φ. Faktorizačním lesem potom budeme rozumět předpis, který pro každé slovo w délky alespoň dva určí, jak jej rozložit na faktory, jinými slovy, udá posloupnost sestávající z alespoň dvou slov, jejichž zřetězením vznikne w. Opakovaným použitím tohoto předpisu potom pro každé slovo w můžeme získat strom jeho faktorů, který vyjadřuje postupný rozklad slova w na písmena.

Definice 2.85. Nechť S je konečná pologrupa a $\varphi\colon A^+ \to S$ je libovolný homomorfismus. *Faktorizační les* pro homomorfismus φ je zobrazení d, které každému slovu nad A délky alespoň 2 přiřadí konečnou posloupnost neprázdných slov nad A délky alespoň 2, tj. $d\colon A^{\geq 2} \to (A^+)^{\geq 2}$, a které splňuje, že pokud $d(w) = (w_1, \dots, w_n)$ pro nějaká slova $w, w_1, \dots, w_n \in A^+$, potom

1. platí $w = w_1 \dots w_n$, tedy d určuje rozklad každého slova na nějaký součin jeho faktorů;

2. pokud $n \geq 3$, tak $\varphi(w)$ je idempotentním prvkem S a je splněno

$$\varphi(w) = \varphi(w_1) = \dots = \varphi(w_n).$$

Faktorizační les tedy pro každé slovo w určuje strom, jehož kořen je ohodnocený w, všechny uzly jsou ohodnoceny nějakými jeho faktory a listy jsou ohodnoceny písmeny. Přitom slovo, kterým je ohodnocen libovolný uzel, získáme zřetězením ohodnocení jeho následníků. Navíc větvení na více než dva následníky je povoleno pouze tehdy, když se všechna slova, jimiž jsou následníci ohodnoceni, vyhodnotí na tentýž idempotentní prvek, a tedy se pro zjištění hodnoty jejich součinu stačí podívat na jeden z nich.

Mezi homomorfismy do konečné pologrupy S hraje význačnou roli *vyhodnocovací homomorfismus* $\mathrm{eval}(S)\colon S^+ \to S$ z volné pologrupy nad prvky pologrupy S, který každý generátor $s \in S$ pologrupy S^+ zobrazuje na příslušný prvek s pologrupy S. Všimněte si, že faktorizační les pro libovolný homomorfismus $\varphi\colon A^+ \to S$ lze získat z faktorizačního lesa pro homomorfismus $\mathrm{eval}(S)$ tak, že slovo $a_1 \dots a_n$, kde $a_1, \dots, a_n \in A$, rozložíme stejně jako slovo $\varphi(a_1) \dots \varphi(a_n)$.

Výškou faktorizačního lesa d rozumíme výšku nejvyššího z jeho stromů. Formálně je výška definována pro každý strom indukcí vzhledem k délce slova:

1. pro $a \in A$ definujeme $h(a) = 0$;

2. pokud $d(w) = (w_1, \dots, w_n)$, tak $h(w) = \max\{h(w_1), \dots, h(w_n)\} + 1$.

Poté položíme $h(d) = \sup\{h(w) \mid w \in A^+\}$.

Cvičení 2.86. Nechť $S = \mathbb{Z}_2$ je dvouprvková grupa a homomorfismus $\varphi\colon \{a,b\}^+ \to S$ je dán předpisem $\varphi(a) = [1]_2$, $\varphi(b) = [0]_2$. Dokažte, že následující předpis definuje faktorizační les pro φ výšky 5.

Je-li $w \in A^{\geq 2}$, potom

1. jestliže je počet výskytů písmene a v w lichý, tak zapíšeme w ve tvaru $w = b^k a\hat{w}$ a položíme
$$d(w) = \begin{cases} (b^k, a), & \text{pokud } \hat{w} = \varepsilon, \\ (b^k a, \hat{w}), & \text{pokud } \hat{w} \neq \varepsilon; \end{cases}$$
2. jestliže je počet výskytů písmene a v w sudý, tak zapíšeme w ve tvaru $w = b^{k_0} a b^{k_1} \dots a b^{k_n}$, kde $k_1, \dots, k_{n-1} \in \mathbb{N}$, $k_0, k_n \in \mathbb{N}_0$, a položíme
$$d(w) = \begin{cases} (a, b^{k_1} a), & \text{pokud } n = 2 \text{ a } k_0 = k_2 = 0, \\ (\underbrace{b, \dots, b}_{k_0}, a b^{k_1} a, \underbrace{b, \dots, b}_{k_2}, \dots, a b^{k_{n-1}} a, \underbrace{b, \dots, b}_{k_n}), & \text{jinak.} \end{cases}$$

Ukažte, že tato výška faktorizačního lesa pro φ je nejmenší možná, neboť každý faktorizační strom slova $a(bbba)^4$ má výšku alespoň 5.

Věta 2.87 (Imre Simon, 1990; Manfred Kufleitner, 2008)**.** *Pro každý homomorfismus z A^+ do konečné pologrupy S existuje faktorizační les výšky $3|S| - 1$. Přitom pro každé slovo $w \in A^+$ je možné příslušný strom zkonstruovat v čase lineárním vzhledem k délce w. Pro $|S| \geq 2$ je výška $3|S| - 1$ obecně nejnižší možná.*

Cvičení 2.88. Dokažte, že existence nějakého faktorizačního lesa konečné výšky je ekvivalentní následujícímu tvrzení:

Pro každý homomorfismus $\varphi\colon A^+ \to S$ do konečné pologrupy existuje regulární výraz reprezentující jazyk všech slov A^* (ekvivalentně, regulární výrazy reprezentující všechny jazyky $\varphi^{-1}(x)$ pro $x \in S$), v němž je iterace aplikována pouze na takové jazyky L, pro které existuje idempotentní prvek $e \in E(S)$ splňující $\varphi(L) = \{e\}$.

Cvičení 2.89. Mějme pevně daný homomorfismus $\varphi\colon A^+ \to S$ do konečné pologrupy. Navrhněte algoritmus, který pro každé slovo $a_1 \dots a_n \in A^+$ vytvoří v čase lineárním vzhledem k n strom ohodnocený prvky S, pomocí kterého poté může určit pro libovolné indexy $i, j \in \{1, \dots, n\}$ splňující $i \leq j$ hodnotu $\varphi(a_i \dots a_j)$ provedením pouze konstantního počtu násobení. Uvědomte si, že nemůžeme prostě předpočítat všechny součiny, neboť faktorů slova $a_1 \dots a_n$ je kvadraticky mnoho.

Důkaz. Dokážeme pouze, že pro každý homomorfismus $\varphi\colon A^+ \to S$ existuje faktorizační les konečné výšky, přičemž nalezení stromu optimální výšky je založeno na stejné myšlence, jen je technicky náročnější. Budeme postupovat indukcí vzhledem k velikosti pologrupy S.

Nejprve si všimněme, že je-li $\psi\colon S \to T$ libovolný homomorfismus, potom k nalezení faktorizačního lesa konečné výšky pro φ stačí nalézt faktorizační lesy konečné výšky pro $\psi \circ \varphi$ a pro jisté homomorfismy do podpologrup pologrupy S tvaru $\psi^{-1}(e)$, kde $e \in E(T)$. Každý faktorizační les d pro $\psi \circ \varphi$ je totiž současně faktorizačním lesem pro φ, s výjimkou pravidel tvaru $d(w) = (w_1, \dots, w_n)$, kde $n \geq 3$ a

$$\psi\varphi(w) = \psi\varphi(w_1) = \cdots = \psi\varphi(w_n) = e$$

pro nějaký idempotentní prvek $e \in E(T)$. Máme-li ovšem k dispozici faktorizační les d' konečné výšky pro homomorfismus $\sigma\colon \{a_1, \dots, a_n\}^+ \to \psi^{-1}(e)$ zadaný předpisem $\sigma(a_i) = \varphi(w_i)$, můžeme rozklad $d(w) = (w_1, \dots, w_n)$ nahradit stromem, který odpovídá stromu slova $a_1 \dots a_n$ v lese d'.

Pokud tedy pologrupa S není jednoduchá, nulová nebo 0-jednoduchá, můžeme uvážit její kvocient podle nějakého maximálního vlastního ideálu a pomocí předchozího argumentu pro ni vytvořit faktorizační les z faktorizačních lesů menších pologrup. Přesněji, faktorizační les pro S lze takto postupně získat z faktorizačních lesů jejích hlavních faktorů.

Protože pro nulové pologrupy je snadné nalézt faktorizační les výšky 3 a jednoduché pologrupy jsou podpologrupami 0-jednoduchých, zbývá popsat faktorizační lesy pro 0-jednoduché pologrupy S. Poněvadž v 0-jednoduché pologrupě je relace $\mathscr{H}$ kongruencí, podle úvodního pozorování stačí důkaz provést pro pologrupu $S/\mathscr{H}$, tedy Reesovu maticovou pologrupu $\mathfrak{M}^0(R, L, \{1\}, P)$, a pro $\mathscr{H}$-třídy obsahující idempotentní prvky, tedy pro grupy.

V případě pologrupy $\mathfrak{M}^0(R, L, \{1\}, P)$, pokud $w \in A^+$ je takové, že $\varphi(w) = 0$, potom uvážíme rozklad w ve tvaru $w = w_1 a_1 \dots w_n a_n w_{n+1}$, kde $w_i \in A^*$ a $a_i \in A$ splňují $\varphi(w_i) \neq 0$ a $\varphi(w_i a_i) = 0$, a položíme

$$\begin{aligned} d(w) &= (w_1 a_1 \dots w_n a_n, w_{n+1}), \\ d(w_1 a_1 \dots w_n a_n) &= (w_1 a_1, \dots, w_n a_n), \\ d(w_i a_i) &= (w_i, a_i). \end{aligned}$$

Tak snadno získáme strom pro w, budeme-li znát stromy pro slova $w_1, \dots, w_{n+1}$. Nyní zbývá zkonstruovat stromy pro slova $w \in A^+$ taková, že $\varphi(w) \neq 0$, přičemž můžeme φ chápat jako homomorfismus do rektangulárního bandu $S = R \times L$. Postupujeme indukcí vzhledem k počtu různých dvojic $(\varphi(a), \varphi(b)) \in S^2$, kde $a, b \in A$ jsou taková, že ab je faktor slova w. V každém kroku zvolíme nějakou dvojici $(x, y) \in S^2$ a rozložíme w do tvaru $w = w_1 a_1 b_1 w_2 \dots w_n a_n b_n w_{n+1}$, kde $a_i b_i$ jsou všechny faktory w délky 2 splňující $(\varphi(a_i), \varphi(b_i)) = (x, y)$ (přitom v případě, že $x = y$, připouštíme i možnost, že některé z faktorů $b_i w_{i+1} a_{i+1}$ slova w mají ve skutečnosti délku jedna a sestávají tedy pouze

z jediného výskytu písmene $b_i = a_{i+1}$). Potom položíme

$$\begin{aligned} d(w) &= (w_1a_1, b_1w_2\dots w_na_nb_nw_{n+1}),\\ d(b_1w_2\dots w_na_nb_nw_{n+1}) &= (b_1w_2\dots w_na_n, b_nw_{n+1}),\\ d(b_1w_2\dots w_na_n) &= (b_1w_2a_2,\dots,b_{n-1}w_na_n) \end{aligned}$$

a využijeme stromy pro slova w_1a_1, $b_1w_2a_2$, ..., $b_{n-1}w_na_n$, b_nw_{n+1}.

Konečně, je-li S grupa, postupujeme indukcí vzhledem k počtu různých prvků $\varphi(v) \in S$, kde v je vlastní prefix slova w. V každém kroku zvolíme nějaký prvek $x \in S$ a rozložíme w do tvaru $w = w_1w_2\dots w_nw_{n+1}$, kde $w_1\dots w_i$, pro $i \in \{1,\dots,n\}$, jsou všechny vlastní prefixy w splňující $\varphi(w_1\dots w_i) = x$. Potom můžeme položit

$$\begin{aligned} d(w) &= (w_1, w_2\dots w_nw_{n+1}),\\ d(w_2\dots w_nw_{n+1}) &= (w_2\dots w_n, w_{n+1}),\\ d(w_2\dots w_n) &= (w_2,\dots,w_n), \end{aligned}$$

neboť $\varphi(w_2) = \dots = \varphi(w_n) = 1$, a využít stromy pro slova w_1, ..., w_{n+1}. Přitom každé ze slov w_i má skutečně méně obrazů vlastních prefixů než w, protože obraz libovolného prefixu v slova w_i lze získat z obrazu prefixu slova w jako $\varphi(v) = \varphi(w_1\dots w_{i-1})^{-1} \cdot \varphi(w_1\dots w_{i-1}v) = x^{-1} \cdot \varphi(w_1\dots w_{i-1}v)$. □

Pomocí faktorizačních lesů se například podařilo najít algoritmus rozhodující problém omezenosti automatů nad polookruhem $(\mathbb{N} \cup \{0,\infty\}, \min, +)$, který jsme zmínili v části 2.5.2. My si jako aplikaci ukážeme větu, která umožňuje algoritmicky rozpoznat, zda je daný regulární jazyk možné vyjádřit v jistém speciálním tvaru, kterému budeme říkat polynom (ve skutečnosti se jedná o jeden případ významné konstrukce, kterou si představíme v části 3.4).

Jazykům tvaru $A_0^*a_1A_1^*\cdots a_kA_k^*$, kde $k \in \mathbb{N}_0$, $a_i \in A$ a $A_i \subseteq A$, říkáme *monomy* stupně k nad abecedou A. *Polynomem* rozumíme konečné sjednocení monomů, přičemž stupeň polynomu definujeme jako nejvyšší stupeň monomu použitého v tomto sjednocení.

Cvičení 2.90. Dokažte, že zřetězením polynomů stupně k a stupně ℓ vznikne polynom stupně nejvýše $k+\ell+1$.

Následující věta využívá důležitý pojem *lokálního podmonoidu* pologrupy S. Tento podmonoid je pro každý idempotentní prvek $e \in E(S)$ definován jako $e \cdot S \cdot e$.

Cvičení 2.91. Dokažte, že pro libovolný idempotentní prvek $e \in E(S)$ je $e \cdot S \cdot e$ vzhledem k inkluzi největší podmonoid pologrupy S, jehož je e neutrálním prvkem.

Cvičení 2.92. Dokažte, že podmonoid $e \cdot S \cdot e$ je průnikem levého a pravého ideálu generovaného prvkem e.

Poznámka 2.93. Konstrukci lokálního podmonoidu lze zobecnit z idempotentních prvků na libovolné prvky $c \in S$ tak, že se na množině $cS^1 \cap S^1c$ definuje operace předpisem $sc \bullet ct = sct$. Takto získáme pojem *lokálního divisoru*, který se podařilo využít k induktivnímu důkazu mnoha důležitých tvrzení.

Věta 2.94 (Jean-Éric Pin a Pascal Weil, 1997)**.** *Pro libovolný regulární jazyk L nad abecedou A jsou následující podmínky ekvivalentní:*

1. *jazyk L je polynomem;*
2. *jazyk L je rozpoznávaný konečným uspořádaným monoidem $(S, \leq)$, v němž je každý idempotentní prvek $e \in E(S)$ nejmenším prvkem lokálního podmonoidu*

$$e \cdot \{x \in S \mid e \leq_{\mathscr{J}} x\}^* \cdot e$$

   v podmonoidu generovaném množinou $\{x \in S \mid e \leq_{\mathscr{J}} x\}$;

3. *pro všechna slova $v, w \in A^*$ splňující $\varphi_L(w) = \varphi_L(w^2)$ a $\mathrm{alph}(v) \subseteq \mathrm{alph}(w)$ platí $w \leq_L wvw$.*

Všimněte si, že poslední podmínku je možné algoritmicky ověřovat.

Důkaz. 1 $\Longrightarrow$ 3: Nechť L je polynom a slova $v, w \in A^*$ splňují předpoklady podmínky 3. Abychom ověřili platnost $w \leq_L wvw$, vezmeme libovolný kontext $(u_1, u_2) \in \mathrm{C}_L(w)$ slova w v jazyce L a dokážeme, že $(u_1, u_2) \in \mathrm{C}_L(wvw)$. Označme n nejvyšší stupeň monomu použitý při konstrukci L. Protože prvek $\varphi_L(w) \in S_L$ je idempotentní, platí $\mathrm{C}_L(w) = \mathrm{C}_L(w^{n+1})$. Proto slovo $u_1 w^{n+1} u_2$ patří do některého z monomů obsažených v L, řekněme

$$u_1 w^{n+1} u_2 \in A_0^* a_1 A_1^* \cdots a_k A_k^* \subseteq L.$$

Jelikož $k \leq n$, leží některá z kopií slova w celá uvnitř některého z výrazů A_i^*. Z toho plyne, že $\mathrm{alph}(v) \subseteq \mathrm{alph}(w) \subseteq A_i$, a tedy přidání slova v na libovolnou stranu vedle této kopie w neovlivní příslušnost slova $u_1 w^{n+1} u_2$ do L. Dostáváme tedy $u_1 w^j v w^k u_2 \in L$ pro jistá $j, k \in \mathbb{N}$. Poněvadž $\mathrm{C}_L(w) = \mathrm{C}_L(w^j) = \mathrm{C}_L(w^k)$, patří i slovo $u_1 wvwu_2$ do L. Proto $(u_1, u_2) \in \mathrm{C}_L(wvw)$.

3 $\Longrightarrow$ 2: Ověříme, že podmínka 2 bude splněna, když k rozpoznání jazyka L použijeme jeho syntaktický homomorfismus $\varphi_L \colon A^* \to (M_L, \leq)$. Buď e libovolný idempotentní prvek monoidu M_L. Označme B množinu všech písmen $a \in A$ splňujících podmínku $\varphi_L(a) \geq_{\mathscr{J}} e$. Protože pro všechna $a \in B$ existují slova $u_a, v_a \in A^*$ taková, že $\varphi_L(u_a a v_a) = e$, můžeme prvek e psát jako $e = \varphi_L(w)$, kde w je slovo vzniklé zřetězením všech slov $u_a a v_a$ pro $a \in B$. Nechť y je libovolný prvek podmonoidu $\{x \in M_L \mid e \leq_{\mathscr{J}} x\}^*$. Jelikož je tento podmonoid generovaný prvky $\varphi_L(a)$, pro $a \in B$, existuje slovo $v \in B^*$ takové, že $y = \varphi_L(v)$. Potom z podmínky 3 dostáváme $e = \varphi_L(w) \leq \varphi_L(wvw) = eye$.

2 $\implies$ 1: Nechť $\varphi\colon A^+ \to (S,\leq)$ je homomorfismus do monoidu splňujícího podmínku 2. Každý jazyk rozpoznávaný tímto homomorfismem je konečným sjednocením jazyků tvaru $\{w \in A^+ \mid \varphi(w) \geq y\}$, pro nějaký prvek $y \in S$. Stačí nám tedy ukázat, že pro daný prvek $y \in S$ je jazyk $\{w \in A^+ \mid \varphi(w) \geq y\}$ polynomem. Za tímto účelem uvážíme faktorizační les d pro φ konečné výšky h, jehož existence je zaručena větou 2.87. Nyní induktivně definujeme pro každé slovo $w \in A^+$ polynom $P(w)$ obsahující w:

$$\begin{aligned}
P(a) &= \{a\}, && \text{pro } a \in A,\\
P(w) &= P(w_1)\cdot P(w_2), && \text{pokud } d(w) = (w_1,w_2),\\
P(w) &= P(w_1)\cdot \mathrm{alph}(w)^*\cdot P(w_n), && \text{pokud } d(w) = (w_1,\ldots,w_n), \text{ kde } n \geq 3.
\end{aligned}$$

Dokážeme, že

$$\{w \in A^+ \mid \varphi(w) \geq y\} = \bigcup \{P(w) \mid w \in A^+,\ \varphi(w) \geq y\},$$

což je polynom, neboť stupeň každého z polynomů $P(w)$ je nejvýše $3\cdot 2^h - 2$ podle cvičení 2.90.

Platnost přímé inkluze vyplývá přímo z faktu $w \in P(w)$. Opačnou inkluzi ověříme tak, že indukcí vzhledem k délce slova w dokážeme, že pro libovolné $v \in P(w)$ platí $\varphi(v) \geq \varphi(w)$. Pro $w \in A$ toto triviálně platí. Jestliže $d(w) = (w_1,w_2)$, tak předpoklad $v \in P(w)$ znamená, že $v = v_1v_2$ pro jistá slova $v_1 \in P(w_1)$ a $v_2 \in P(w_2)$. Podle indukčního předpokladu platí $\varphi(v_1) \geq \varphi(w_1)$ a $\varphi(v_2) \geq \varphi(w_2)$. Proto

$$\varphi(v) = \varphi(v_1v_2) \geq \varphi(w_1w_2) = \varphi(w).$$

Konečně, je-li $d(w) = (w_1,\ldots,w_n)$, kde $n \geq 3$, tak $v = v_1uv_n$ pro jistá $v_1 \in P(w_1)$, $u \in \mathrm{alph}(w)^*$ a $v_n \in P(w_n)$. Z indukčního předpokladu dostáváme $\varphi(v_1) \geq \varphi(w_1)$ a $\varphi(v_n) \geq \varphi(w_n)$. Dále platí

$$\varphi(u) \in \{x \in S \mid \varphi(w) \leq_{\mathscr{J}} x\}^*,$$

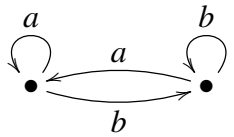
což ověříme tak, že za jednotlivé prvky x bereme obrazy písmen slova u. Protože z definice faktorizačního lesa plyne, že $\varphi(w_1) = \varphi(w_n) = \varphi(w)$ je idempotentní prvek S, díky předpokládané vlastnosti monoidu $(S,\leq)$ dostáváme

$$\varphi(v) = \varphi(v_1)\varphi(u)\varphi(v_n) \geq \varphi(w_1)\varphi(u)\varphi(w_n) = \varphi(w)\varphi(u)\varphi(w) \geq \varphi(w). \qquad \square$$

2.7 Věnčité a kaskádové součiny

V této části si ukážeme nejdůležitější konstrukci, která z daných dvou pologrup vytváří pologrupu složitější. Jedná se o věnčitý součin a jeho význam spočívá v tom, že umožňuje každou konečnou pologrupu rozložit na součin velmi jednoduchých nerozložitelných pologrup. Těmito nerozložitelnými pologrupami jsou jednak jednoduché grupy

a jednak podpologrupy monoidu U_2, zvaného také *flip-flop monoid*, což je tříprvkový monoid obsahující dva pravé nulové prvky. Název tohoto monoidu vychází z faktu, že je antiizomorfní přechodovému monoidu automatu



Na druhou stranu odpovídá ovšem věnčitý součin přirozené konstrukci s jazyky, a proto se hojně používá k získání efektivních charakterizací tříd regulárních jazyků.

Dosud jsme se setkali s jedinou operací, která umožňuje vytvářet monoidy, respektive automaty, pro rozpoznávání nových jazyků, a to s přímým součinem. Jak ale ukazuje následující cvičení, možnosti přímého součinu jsou velmi omezené.

Cvičení 2.95. Nechť $\mathscr{A} = (A^*, Q, \delta, q_0)$ a $\mathscr{B} = (A^*, P, \rho, p_0)$ jsou dva automaty bez vyznačených koncových stavů. Dokažte, že provedeme-li v jejich přímém součinu $\mathscr{A} \times \mathscr{B} = (A^*, Q \times P, \delta \times \rho, (q_0, p_0))$ všechny možné volby koncových stavů, dostaneme automaty přijímající právě jazyky, které jsou booleovskými kombinacemi jazyků přijímaných automaty $\mathscr{A}$ a $\mathscr{B}$ při různých volbách koncových stavů. Uvědomte si, že jako speciální případ právě dokázaného faktu dostáváme pro libovolnou dvojici homomorfismů $\varphi\colon A^* \to S$ a $\psi\colon A^* \to T$ charakterizaci jazyků rozpoznávaných homomorfismem $(\varphi, \psi)\colon A^* \to S \times T$.

Nejprve si popíšeme konstrukci na deterministických automatech odpovídající věnčitému součinu, která se nazývá kaskádový součin. V tomto součinu čtou vstup oba automaty současně jako v případě přímého součinu, ovšem akce prvního automatu při čtení každého písmene závisí i na tom, v jakém stavu se právě nachází druhý automat. Protože nás nyní nebude zajímat přímo jazyk přijímaný daným automatem, ale to, jaké jazyky může tento automat přijímat při různé volbě počátečních a koncových stavů, budeme místo s automaty pracovat s akcemi na množině stavů.

Definice 2.96. *Kaskádový součin* akcí $\delta\colon Q \times A \to Q$ nad abecedou A a $\rho\colon P \times (Q \times A) \to P$ nad abecedou $Q \times A$ je akce $\rho \circ \delta\colon (P \times Q) \times A \to P \times Q$ nad abecedou A definovaná na množině stavů $P \times Q$ předpisem

$$(\rho \circ \delta)((p,q),a) = (\rho(p,(q,a)), \delta(q,a)). \tag{2.1}$$

Všimněte si, že akce $\rho \circ \delta$ působí na libovolném slově $w = a_1 \dots a_n \in A^+$ předpisem

$$(\rho \circ \delta)((p,q),w) = (\rho(p,(q,a_1)(\delta(q,a_1),a_2)\dots(\delta(q,a_1\dots a_{n-1}),a_n)), \delta(q,w)). \tag{2.2}$$

Následující cvičení ukazuje, že (považujeme-li kartézský součin množin za asociativní) je kaskádový součin asociativní operace.

Cvičení 2.97. Dokažte, že pro libovolné akce $\delta\colon Q \times A \to Q$, $\rho\colon P \times (Q \times A) \to P$ a $\tau\colon R \times (P \times Q \times A) \to R$ jsou akce $(\tau \circ \rho) \circ \delta$ a $\tau \circ (\rho \circ \delta)$ izomorfní prostřednictvím přirozené bijekce mezi množinami stavů $(R \times P) \times Q$ a $R \times (P \times Q)$.

Teď si ukážeme, že kaskádový součin $\rho \circ \delta$ dokáže přijímat vzory jazyků přijímaných akcí ρ v jisté racionální relaci $\sigma \subseteq A^* \times (Q \times A)^*$. Uvažme konečný převodník, který má za množinu stavů Q a jehož hrany jsou právě $(q, (a, (q, a)), \delta(q, a))$, přičemž počátečním stavem je libovolný $q_0 \in Q$ a množinou koncových stavů je libovolná $F \subseteq Q$. Přechody tohoto převodníku jsou tedy stejné jako přechody akce δ, přičemž výstupem každé hrany je dvojice (q, a) obsahující informaci jak o vstupním písmenu, tak o aktuálním stavu. Relace σ počítaná tímto převodníkem je parciální funkce, kterou lze při označení

$$\varsigma(a_1 \dots a_n) = (q_0, a_1)(\delta(q_0, a_1), a_2) \dots (\delta(q_0, a_1 \dots a_{n-1}), a_n), \tag{2.3}$$

zadat předpisem

$$\sigma(w) = \begin{cases} \varsigma(w), & \text{jestliže } \delta(q_0, w) \in F, \\ \text{nedefinováno}, & \text{jinak.} \end{cases} \tag{2.4}$$

Přepíšeme-li vzorec (2.2) pomocí označení (2.3), dostaneme

$$(\rho \circ \delta)((p_0, q_0), w) = (\rho(p_0, \varsigma(w)), \delta(q_0, w)).$$

Výpočet na první složce kaskádového součinu $\rho \circ \delta$ tedy odpovídá výpočtu akce ρ na slově $\sigma(w)$, přičemž druhá složka kaskádového součinu rozhoduje o tom, zda je parciální funkce σ na slově w opravdu definována. Proto tento součin umí na dané slovo nad abecedou A aplikovat funkci σ a poté testovat příslušnost výsledného slova do nějakého jazyka přijímaného akcí ρ. Přesněji, automat $\mathscr{A} = (A^*, P \times Q, \rho \circ \delta, (p_0, q_0), H \times F)$ přijímá jazyk $\sigma^{-1}(L(\mathscr{B}))$, kde $\mathscr{B}$ značí automat $((Q \times A)^*, P, \rho, p_0, H)$.

Podívejme se nyní, jaké přechody mezi stavy v P a mezi stavy v Q určují působení písmene $a \in A$ na kaskádovém součinu. Na množině stavů Q působí a transformací $t_a \in \mathscr{T}(Q)$ danou předpisem $t_a(q) = \delta(q, a)$. Na množině stavů P působí a pro každé $q \in Q$ jednou transformací. Proto tomuto písmenu odpovídá zobrazení $f_a\colon Q \to \mathscr{T}(P)$ dané předpisem $f_a(q)(p) = \rho(p, (q, a))$. Definiční předpis kaskádového součinu (2.1) můžeme potom přepsat jako

$$(\rho \circ \delta)((p, q), a) = (f_a(q)(p), t_a(q)). \tag{2.5}$$

Uvažme dále nějaké písmeno $b \in A$, které stejným způsobem určuje zobrazení $t_b \in \mathscr{T}(Q)$ a $f_b \in \mathscr{T}(P)^Q$. Vyjádřeme nyní působení slova ab na kaskádovém součinu $\rho \circ \delta$ pomocí zobrazení f_a, t_a, f_b a t_b:

$$(\rho \circ \delta)((p, q), ab) = \Big(\big(f_b(t_a(q)) \circ f_a(q)\big)(p), (t_b \circ t_a)(q) \Big). \tag{2.6}$$

Označíme-li $t_{ab} \in \mathscr{T}(Q)$ a $f_{ab} \in \mathscr{T}(P)^Q$ zobrazení vyjadřující působení slova ab na $\rho \circ \delta$, analogicky (2.5) máme

$$(\rho \circ \delta)((p,q),ab) = (f_{ab}(q)(p), t_{ab}(q)).$$

Proto $t_{ab} = t_b \circ t_a$ a pro všechna $q \in Q$ je

$$f_{ab}(q) = f_b(t_a(q)) \circ f_a(q). \tag{2.7}$$

Tento předpis nyní vyjádříme pomocí jisté akce φ monoidu $\mathscr{T}(Q)$ na monoidu $\mathscr{T}(P)^Q$, který chápeme jako mocninu monoidu $\mathscr{T}(P)$, tedy s operacemi po složkách. Akci φ definujeme pro libovolná $t \in \mathscr{T}(Q)$, $f \in \mathscr{T}(P)^Q$ a $q \in Q$ předpisem

$$\varphi(t)(f)(q) = f(t(q)).$$

Formuli (2.7) potom můžeme přepsat jako

$$f_{ab} = \varphi(t_a)(f_b) \cdot f_a,$$

kde $\cdot$ je operace monoidu $\mathscr{T}(P)^Q$. Dohromady získáme následující předpis pro násobení dvojic zobrazení vyjadřujících chování slov na $\rho \circ \delta$:

$$(f_a, t_a) \cdot (f_b, t_b) = (\varphi(t_a)(f_b) \cdot f_a, t_b \circ t_a).$$

Abstraktní definice věnčitého součinu monoidů je založena na této formuli v případě, kdy za množinu Q zvolíme nějaký monoid T a za δ akci T na sobě pravým násobením, tedy $\varphi(t)(f)(q) = f(q \cdot t)$. Formální definice věnčitého součinu využívá pojmu polopřímého součinu, který je zobecněním přímého součinu na případ, kdy prvky na druhé složce pomocí nějaké akce ovlivňují výsledek násobení na první složce. Přitom *akcí* pologrupy T na pologrupě S myslíme libovolný homomorfismus $\varphi\colon T \to \mathrm{End}(S)$, kde $\mathrm{End}(S)$ je monoid všech homomorfismů pologrupy S do sebe. (Na rozdíl od automatů definujeme tentokrát akci jako homomorfismus, neboť bývá zvykem používat předpis pro polopřímý součin, který vyžaduje aplikaci akce zleva.)

Definice 2.98. *Polopřímý součin* $S \rtimes_\varphi T$ pologrup S a T vzhledem k akci $\varphi\colon T \to \mathrm{End}(S)$ je definován na množině $S \times T$ předpisem

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 \cdot \varphi(y_1)(x_2), y_1 \cdot y_2).$$

Všimněte si, že zvolíme-li akci φ triviální, tedy $\varphi(y)(x) = x$ pro všechna $y \in T$ a $x \in S$, dostaneme definici přímého součinu.

Cvičení 2.99. Dokažte, že monoid všech afinních transformací konečněrozměrného afinního prostoru A se zaměřením V je izomorfní polopřímému součinu aditivní grupy vektorového prostoru V a monoidu lineárních transformací prostoru V, vzhledem k přirozené akci lineárních transformací na prostoru V.

Cvičení 2.100. Dokažte, že předpis v předchozí definici opravdu zadává asociativní operaci na $S \times T$. Dále ukažte, že pokud S a T jsou grupy, potom $S \rtimes_\varphi T$ je rovněž grupa, pokud S a T jsou aperiodické pologrupy, potom pologrupa $S \rtimes_\varphi T$ je rovněž aperiodická, a pokud mají pologrupy S a T Greenovu relaci $\mathscr{R}$ triviální, potom i pologrupa $S \rtimes_\varphi T$ má $\mathscr{R}$ triviální.

Věnčitý součin je speciálním případem polopřímého součinu, kde první pologrupa má určitou strukturu, která umožňuje druhé pologrupě na ní přirozeně působit pomocí pravého násobení. Protože definice věnčitého součinu pro pologrupy není ustálená, formulujeme následující definici pouze pro monoidy, což je pro prezentaci rozkladu pologrup postačující. Pro libovolné monoidy S a T budeme uvažovat monoid S^T, který je přímým součinem kopií S, což znamená, že násobení se na něm provádí po složkách: pro $f, g \in S^T$ máme $(f \cdot g)(t) = f(t) \cdot g(t)$. Nyní definujeme akci φ monoidu T na monoidu S^T předpisem

$$\varphi(t')(f)(t) = f(tt'), \tag{2.8}$$

pro $t, t' \in T$ a $f \in S^T$. Násobení prvky T zprava je tedy použito k prohazování jednotlivých souřadnic v součinu S^T.

Cvičení 2.101. Dokažte, že předcházející předpis skutečně definuje akci monoidu T na monoidu S^T.

Definice 2.102. *Věnčitý součin* (wreath product) $S \wr T$ monoidů S a T je definován jako polopřímý součin $S^T \rtimes_\varphi T$, kde akce φ je určena předpisem (2.8).

Uvědomte si, že věnčitý součin $S \wr T$ je vytvořen z monoidů S a T pouze použitím polopřímých součinů, a proto má operace věnčitého součinu všechny uzávěrové vlastnosti uvedené ve cvičení 2.100.

Cvičení 2.103. Dokažte, že pro libovolné monoidy S a T a pro libovolnou akci $\varphi\colon T \to \mathrm{End}(S)$ je polopřímý součin $S \rtimes_\varphi T$ izomorfní podmonoidu věnčitého součinu $S \wr T$, přičemž příslušné vnoření je dáno předpisem $\psi(s,t) = (f,t)$, kde $f(u) = \varphi(u)(s)$ pro všechna $u \in T$.

Všimněte si, že věnčitý součin nemůže být asociativní operací na třídě konečných monoidů, již kvůli různým mohutnostem výsledných nosičů. Platí ovšem následující slabší tvrzení:

Cvičení 2.104. Dokažte, že pro libovolné monoidy S, T, U je součin $(S \wr T) \wr U$ izomorfní podmonoidu součinu $S \wr (T \wr U)$, přičemž toto vnoření je dáno předpisem

$$\begin{aligned} \psi\colon (S^T)^U \times T^U \times U &\to S^{T^U \times U} \times T^U \times U \\ \psi(f,g,u) &= (h,g,u), \end{aligned}$$

kde

$$h(k,v) = f(v)(k(1)),$$

pro všechna $k \in T^U$ a $v \in U$.

Iterovaný věnčitý součin $S_1 \wr \cdots \wr S_n$ tedy definujeme jako $S_1 \wr (S_2 \wr \cdots \wr (S_{n-1} \wr S_n) \cdots)$, aby byl výsledný monoid co možná největší.

Cvičení 2.105. *** vzorec pro iterovany vencity soucin v jednodussim poradi

Cvičení 2.106. Dokažte, že pokud monoid U dělí monoid S, tak monoid $U \wr T$ dělí monoid $S \wr T$.

Použití věnčitých součinů v teorii regulárních jazyků si budeme ilustrovat na *$\mathscr{R}$-triviálních* monoidech, tedy monoidech, jejichž všechny $\mathscr{R}$-třídy jsou jednoprvkové. Nejprve si ukážeme, jak je možné tyto monoidy pomocí věnčitého součinu rozložit. Základním stavebním kamenem přitom bude monoid U_1, což je dvouprvkový monoid sestávající z nulového a neutrálního prvku.

Lemma 2.107. *Konečná pologrupa je $\mathscr{R}$-triviální právě tehdy, když dělí iterovaný věnčitý součin kopií monoidu U_1.*

Důkaz. *** □

K rozkladu libovolné konečné grupy lze použít známou větu, která říká, že věnčitý součin dvou grup $H \wr K$ vždy obsahuje (až na izomorfismus) všechna rozšíření grupy H pomocí grupy K; připomeňme, že rozšířením grupy H pomocí grupy K se nazývá libovolná grupa G, která obsahuje H jako normální podgrupu, přičemž příslušný kvocient G/H je izomorfní K. Všimněte si, že libovolný polopřímý součin $H \rtimes_\varphi K$ je izomorfní nějakému rozšíření grupy H pomocí grupy K.

Věta 2.108 (Marc Krasner a Lev Arkaděvič Kalužnin, 1951)**.** *Je-li H normální podgrupa grupy G, potom G je izomorfní podgrupě věnčitého součinu $H \wr (G/H)$.*

Abychom tedy konečnou grupu G rozložili jako věnčitý součin jednoduchých grup, stačí vzít libovolnou kompoziční řadu grupy G, to znamená posloupnost podgrup

$$\{1\} \lhd H_1 \lhd H_2 \lhd \cdots \lhd H_n = G,$$

kde každá podgrupa H_i je maximální normální podgrupou H_{i+1}, tedy každý kvocient H_{i+1}/H_i je jednoduchá grupa. Potom totiž grupa G podle předchozí věty a cvičení 2.106 dělí součin jednoduchých grup

$$(\cdots(H_1 \wr (H_2/H_1)) \wr \cdots \wr (H_{n-1}/H_{n-2})) \wr (G/H_{n-1}).$$

Mnohem obtížnější ovšem je ukázat následující tvrzení pro obecné pologrupy:

Věta 2.109 (Kenneth Krohn a John L. Rhodes, 1965)**.** *Libovolná konečná pologrupa S dělí nějaký iterovaný věnčitý součin monoidu U_2 a jednoduchých grup dělících S.*

Poznámka 2.110. Většinou se věta 2.109 formuluje obecněji pro věnčitý součin konečných transformačních monoidů, tedy akcí konečného monoidu na konečné množině.

Předchozí větu je samozřejmě možné formulovat i pro kaskádový součin automatů. Z tohoto pohledu potom říká, že každý konečný automat je možné hierarchicky rozložit takovým způsobem, že informace se předávají mezi komponentami rozkladu pouze jedním směrem, přičemž každá komponenta je buď čítač v podobě konečné jednoduché grupy nebo automat resetující se do jednoho ze dvou stavů.

Jazyky rozpoznávané věnčitým součinem dvou monoidů dosti přesně popisuje následující tvrzení, které bývá často nazýváno *princip věnčitého součinu*. Pro jeho formulaci je klíčový pojem *sekvenčního převodníku*, což je konečný převodník deterministický v první složce, tedy z každého stavu vede pro každé písmeno vstupní abecedy právě jedna hrana taková, že první složka jejího ohodnocení je rovna tomuto písmenu, a jiné hrany tento převodník neobsahuje. Je jasné, že relace přijímaná takovým převodníkem je vždy parciální funkce, a takové parciální funkci se říká *sekvenční funkce*. Pro libovolný sekvenční převodník můžeme mluvit o jeho přechodovém monoidu, čímž myslíme přechodový monoid deterministického automatu, který z něj vznikne odstraněním výstupních ohodnocení. Typickým příkladem takového převodníku je výše popsaný převodník přijímající funkci σ danou předpisem (2.4). Takto definované převodníky jsou jádrem důkazu avizovaného tvrzení.

Věta 2.111 (Howard Straubing, 1979)**.** *Nechť S a T jsou konečné monoidy.*

Je-li σ sekvenční funkce z A^* do B^* přijímaná sekvenčním převodníkem, jehož přechodový monoid je antiizomorfní T, potom pro každý jazyk $L \subseteq B^*$ rozpoznávaný monoidem S je jazyk $\sigma^{-1}(L) \subseteq A^*$ rozpoznávaný monoidem $S \wr T$.

Opačně, každý jazyk rozpoznávaný monoidem $S \wr T$ lze získat pomocí booleovských operací z jazyků tvaru $\sigma^{-1}(L)$, kde L je rozpoznávaný monoidem S a σ je sekvenční funkce přijímaná sekvenčním převodníkem, jehož přechodový monoid je antiizomorfní T.

Důkaz. *** □

Jako příklad aplikace této věty si ukážeme, jaké jazyky rozpoznávají $\mathscr{R}$-triviální monoidy. Nejprve si uvědomme, jakým tyto monoidy odpovídají automatům:

Cvičení 2.112. Dokažte, že syntaktický monoid regulárního jazyka L je $\mathscr{R}$-triviální právě tehdy, když minimální automat jazyka L neobsahuje žádné cykly délky alespoň dva, tedy jestliže pro všechny stavy q a písmena a a b z platnosti $\delta(q,ab) = q$ plyne $\delta(q,a) = q$.

Větu 2.111 použijeme ke zjištění, jaké jazyky rozpoznává věnčitý součin daného monoidu s jednou kopií monoidu $U_1 = \{0,1\}$:

Lemma 2.113. *Nechť T je libovolný konečný monoid. Potom každý jazyk nad abecedou A rozpoznávaný monoidem $U_1 \wr T$ je booleovskou kombinací jazyků rozpoznávaných monoidem T a jazyků tvaru KaA^*, kde K je jazyk rozpoznávaný T a $a \in A$.*

Důkaz. Nechť $\sigma\colon A^* \to B^*$ je libovolná sekvenční funkce přijímaná sekvenčním převodníkem $\mathscr{A}$ s množinou stavů Q, jehož přechodový monoid je antiizomorfní T. Nejprve si uvědomme, že monoidem U_1 jsou rozpoznávané právě jazyky tvaru B^*CB^* nebo C^* pro nějakou abecedu B a její podmnožinu $C \subseteq B$. Každý jazyk tvaru $\sigma^{-1}(L)$, kde L je rozpoznávaný U_1, je tedy booleovskou kombinací jazyků tvaru $\sigma^{-1}(B^*bB^*)$, pro nějaké písmeno $b \in B$. Nechť $\mathscr{A}'$ je automat nad abecedou A vzniklý z převodníku $\mathscr{A}$ odstraněním všech výstupů. Označme $K \subseteq A^*$ jazyk rozpoznávaný automatem $\mathscr{A}'$ a $K_q \subseteq A^*$, pro $q \in Q$, jazyk rozpoznávaný automatem vzniklým z automatu $\mathscr{A}'$ tak, že za jediný koncový stav zvolíme q. Jazyk $\sigma^{-1}(B^*bB^*)$ je potom sjednocením všech jazyků $K_qaA^* \cap K$, pro $q \in Q$ a $a \in A$ takové, že při přechodu písmenem a ze stavu q produkuje $\mathscr{A}$ na výstupu slovo obsahující písmeno b. □

Věta 2.114. *Jazyk nad abecedou A je rozpoznatelný konečným $\mathscr{R}$-triviálním monoidem právě tehdy, když je konečným sjednocením jazyků tvaru*

$$A_0^* a_1 A_1^* \cdots a_k A_k^*, \qquad \text{kde } k \in \mathbb{N}_0,\ A_i \subseteq A \text{ a } a_i \in A \setminus A_{i-1}. \tag{2.9}$$

Všimněte si, že v případě obecných polynomů můžeme písmena $a_1, \ldots, a_k$ vybírat kdekoli v daném slově, zatímco racionální výrazy v předchozí větě vyžadují, abychom vždy použili první vyhovující výskyt daného písmene.

Cvičení 2.115. Ukažte, že jazyk $A^* a_1 A^* \cdots a_k A^*$ je tvaru (2.9). Dejte příklad polynomu, jehož syntaktický monoid není $\mathscr{R}$-triviální.

Cvičení 2.116. Dokažte, že třída jazyků, které jsou konečným sjednocením jazyků (2.9), je uzavřená na komplementy.

Cvičení 2.117. Dokažte větu 2.114 pomocí lemmat 2.107 a 2.113 a cvičení 2.116.

Všimněte si, že podle věty 2.109 každý aperiodický monoid dělí iterovaný součin flip-flop monoidů, což je možné použít k získání alternativního důkazu Schützenbergerho věty 2.71, který je založený na charakterizaci jazyků rozpoznávaných věnčitým součinem $U_2 \wr T$ získané podobně jako v lemmatu 2.113.

Kapitola 3

Variety jazyků

V předchozí kapitole jsme viděli, že mnohé kombinatorické vlastnosti regulárních jazyků je možné charakterizovat pomocí algebraických vlastností jejich syntaktických monoidů. Cílem této kapitoly je prezentovat obecnou teorii popisující vztah mezi jistými třídami regulárních jazyků zvanými variety a pseudovarietami konečných monoidů či pologrup. Tento vztah byl objeven Samuelem Eilenbergem v roce 1976. Celou kapitolu opět pokrývá nová kniha Jean-Érica Pina [6].

V následujícím textu vybudujeme teorii pseudovariet pouze v případě monoidů. Získané výsledky je ovšem snadné rozšířit nejen na případ pologrup, ale dokonce na všechny algebry konečného typu. Případné rozdíly a vztahy mezi teorií pro monoidy a teorií pro pologrupy v textu zmíníme.

3.1 Pseudovariety

Pseudovariety jsou analogií pojmu variet, známého z univerzální algebry, v případě, že studované třídy algeber obsahují pouze konečné algebry. Podobně jako se definují variety algeber jako třídy algeber daného typu uzavřené na homomorfní obrazy, podalgebry a součiny, pseudovariety algeber se definují jako třídy konečných algeber uzavřené na homomorfní obrazy, podalgebry a konečné součiny.

Definice 3.1. Třída $\mathbf{V}$ konečných monoidů se nazývá *pseudovarieta*, jestliže je neprázdná a splňuje následující podmínky:

1. Je-li $\psi\colon S \twoheadrightarrow T$ surjektivní homomorfismus monoidů a platí $S \in \mathbf{V}$, potom platí i $T \in \mathbf{V}$.

2. Je-li T podmonoid monoidu $S \in \mathbf{V}$, potom platí $T \in \mathbf{V}$.

3. Pokud platí $S, T \in \mathbf{V}$, potom platí i $S \times T \in \mathbf{V}$.

Cvičení 3.2. Pro třídu konečných monoidů $\mathbf{T}$ značíme $\mathrm{H}(\mathbf{T})$ třídu všech homomorfních obrazů monoidů z $\mathbf{T}$, $\mathrm{S}(\mathbf{T})$ třídu všech podmonoidů monoidů z $\mathbf{T}$ a $\mathrm{P}_{\mathrm{fin}}(\mathbf{T})$ třídu všech konečných součinů monoidů z $\mathbf{T}$. Dokažte, že pro libovolnou třídu konečných monoidů $\mathbf{T}$ je $\mathrm{HSP}_{\mathrm{fin}}(\mathbf{T})$ pseudovarieta generovaná $\mathbf{T}$.

Definici pseudovariet můžeme analogicky formulovat i pro pologrupy, přičemž místo monoidových homomorfismů uvažujeme homomorfismy pologrup a místo podmonoidů podpologrupy.

Cvičení 3.3. Ukažte, že svaz všech pseudovariet konečných monoidů je přirozeně izomorfní úplnému podsvazu svazu všech pseudovariet konečných pologrup; přitom pseudovarietě monoidů $\mathbf{V}$ přiřadíme pseudovarietu pologrup obsahující právě pologrupy S takové, že $S^1 \in \mathbf{V}$, což je ekvivalentní požadavku, že S je podpologrupou nějakého monoidu z $\mathbf{V}$. Dále ukažte, že ne každá pseudovarieta pologrup lze tímto postupem z nějaké pseudovariety monoidů získat; příkladem takové pseudovariety je pseudovarieta všech rektangulárních bandů.

Podobně jako v případě variet definujme pro každou pseudovarietu $\mathbf{V}$ a pro libovolnou množinu A na volném monoidu A^* kongruenci $\sim_{\mathbf{V}}^{A}$ složenou právě z identit platných v pseudovarietě $\mathbf{V}$, tedy $u \sim_{\mathbf{V}}^{A} v$ platí právě tehdy, když pro každý homomorfismus $\varphi\colon A^* \to S$ do libovolného monoidu $S \in \mathbf{V}$ je splněno $\varphi(u) = \varphi(v)$. Označme $\mathrm{F}_{\mathbf{V}}(A)$ příslušný kvocient $A^*/{\sim_{\mathbf{V}}^{A}}$. Podle Birkhoffovy věty je každá varieta monoidů $\mathscr{V}$ jednoznačně popsaná příslušnou kongruencí $\sim_{\mathscr{V}}^{A}$ na A^* pro spočetnou množinu A: libovolný monoid S totiž patří do $\mathscr{V}$ právě tehdy, když splňuje všechny identity z $\sim_{\mathscr{V}}^{A}$, což nastává právě tehdy, když je homomorfním obrazem monoidu $\mathrm{F}_{\mathscr{V}}(B) = B^*/{\sim_{\mathscr{V}}^{B}}$ pro nějakou množinu B. Hlavním důvodem je, že díky uzavřenosti variet na libovolné součiny patří všechny monoidy $\mathrm{F}_{\mathscr{V}}(B)$ do $\mathscr{V}$. Jelikož pseudovariety obsahují pouze konečné monoidy, ale monoid $\mathrm{F}_{\mathbf{V}}(B)$ je často nekonečný i pro konečnou množinu B, nemůžeme podobnou vlastnost od pseudovariet očekávat.

Cvičení 3.4. Dokažte, že pro libovolnou pseudovarietu monoidů $\mathbf{V}$ a libovolnou konečnou množinu A platí $\mathrm{F}_{\mathbf{V}}(A) \in \mathbf{V}$ právě tehdy, když je monoid $\mathrm{F}_{\mathbf{V}}(A)$ konečný.

Pro libovolnou varietu monoidů $\mathscr{V}$ je zřejmě třída $\mathscr{V}_{\mathrm{fin}}$ všech konečných monoidů patřících do $\mathscr{V}$ pseudovarieta. Podle Birkhoffovy věty tedy takto získáme právě pseudovariety, které lze popsat pomocí identit. Jak dále uvidíme, většina zajímavých pseudovariet ovšem takto získat nelze. Příkladem takové pseudovariety je pseudovarieta všech konečných grup $\mathbf{G}$:

Příklad 3.5. Ukážeme sporem, že pseudovarieta $\mathbf{G}$ nesplňuje žádné netriviální monoidové identity a generuje tedy varietu všech monoidů. Nechť $a_1 \dots a_m = b_1 \dots b_n$ je nejkratší monoidová identita splněná v $\mathbf{G}$, přičemž $a_1, \dots, a_m, b_1, \dots b_n \in A$ a $m \geq n$. Jistě platí $n \geq 1$, neboť v konečných grupách existují prvky libovolného konečného řádu.

Díky minimalitě této identity musí být písmena a_m a b_n různá. Uvažme grupu S_{2m+1} všech permutací množiny $\{1,\dots,2m+1\}$ a dosaďme za všechna písmena a_i a b_i cykly $\rho = (m+1, m, \dots, 1)$ a $\sigma = (m+1, m+2, \dots, 2m+1)$, přičemž za a_m dosadíme ρ a za b_n dosadíme σ. Potom se snadno vidí, že slovo $a_1 \dots a_m$ se vyhodnotí na permutaci, která prvek $m+1$ zobrazuje na některý z prvků m, ..., 1, zatímco slovo $b_1 \dots b_n$ se vyhodnotí na permutaci, která prvek $m+1$ zobrazuje na některý z prvků $m+2$, ..., $2m+1$.

Všimněte si ale, že pseudovarietu **G** je možné popsat pomocí obecnějších rovností než identit, neboť konečný monoid S je grupou právě tehdy, když pro libovolný prvek $x \in S$ platí $x^\omega = 1$.

Jiným příkladem pseudovariety nepopsatelné identitami je pseudovarieta všech konečných nilpotentních pologrup.

Definice 3.6. Říkáme, že pologrupa S je *nilpotentní*, jestliže existuje $n \in \mathbb{N}$ takové, že výsledkem libovolného součinu n prvků pologrupy S je nulový prvek pologrupy S.

Snadno se vidí, že konečné nilpotentní pologrupy tvoří pseudovarietu, kterou značíme **N**.

Cvičení 3.7. Dokažte, že pseudovarieta **N** generuje varietu všech pologrup. Dále dokažte, že konečná pologrupa S je nilpotentní právě tehdy, když pro libovolné prvky $x, y \in S$ platí $x^\omega y = y x^\omega = x^\omega$, přičemž tyto dvě rovnosti bývá zvykem stručně zapisovat dohromady jako $x^\omega = 0$.

Cvičení 3.8. Buď S libovolný konečný monoid a nechť **V** je pseudovarieta generovaná S a $\mathscr{V}$ varieta generovaná S. Dokažte, že platí $\mathbf{V} = \mathscr{V}_{\mathrm{fin}}$.

Vyvoďte z tohoto tvrzení, že každou pseudovarietu generovanou konečně mnoha monoidy je možné zadat pomocí identit.

Tvrzení 3.9. *Pro každou pseudovarietu konečných monoidů* **V** *existuje nekonečná posloupnost variet monoidů* $\mathscr{V}_1 \subseteq \mathscr{V}_2 \subseteq \cdots$ *taková, že* $\mathbf{V} = \bigcup_{n\in\mathbb{N}} (\mathscr{V}_n)_{\mathrm{fin}}$.

Naopak, sjednocením libovolného řetězce pseudovariet je pseudovarieta.

Důkaz. Protože konečných monoidů je až na izomorfismus spočetně mnoho, můžeme všechny monoidy patřící do **V** vyčíslit jako S_n, pro $n \in \mathbb{N}$. Vezmeme-li za $\mathscr{V}_n$ varietu generovanou množinou $\{S_1, \dots, S_n\}$, bude triviálně splněno $\mathbf{V} \subseteq \bigcup_{n\in\mathbb{N}} (\mathscr{V}_n)_{\mathrm{fin}}$ a platnost opačné inkluze, tedy $(\mathscr{V}_n)_{\mathrm{fin}} \subseteq \mathbf{V}$, vyplyne ze cvičení 3.8.

Platnost druhé části tvrzení se snadno ověří. □

Tvrzení 3.10. *Třída* **V** *konečných monoidů je pseudovarieta právě tehdy, když existuje posloupnost identit* ι_k, $k \in \mathbb{N}$, *taková, že* **V** *sestává právě z konečných monoidů, které splňují skoro všechny identity* ι_k.

Důkaz. Podle tvrzení 3.9 existují pseudovariety $\mathbf{V}_n$ definované identitami takové, že $\mathbf{V}_1 \subseteq \mathbf{V}_2 \subseteq \cdots$ a $\mathbf{V} = \bigcup_{n\in\mathbb{N}} \mathbf{V}_n$. Nechť T_m, pro $m \in \mathbb{N}$, jsou až na izomorfismus všechny konečné monoidy nepatřící do $\mathbf{V}$. Zvolme pro všechna $m,n \in \mathbb{N}$ identitu $\iota_{m,n}$, která platí ve $\mathbf{V}_n$, ale není splněná v T_m. Ukážeme, že konečný monoid S patří do $\mathbf{V}$ právě tehdy, když splňuje skoro všechny identity $\iota_{m,n}$ pro $n \geq m$. Pokud platí $S \in \mathbf{V}$, tak existuje $n_0 \in \mathbb{N}$ takové, že $S \in \mathbf{V}_n$ pro všechna $n \geq n_0$. Proto splňuje S všechny identity $\iota_{m,n}$ kromě těch s $n < n_0$, kterých ovšem díky předpokladu $n \geq m$ uvažujeme jen konečně mnoho. Pokud naopak S do $\mathbf{V}$ nepatří, pak je izomorfní některému z monoidů T_m, a tedy nesplňuje žádnou z nekonečně mnoha identit $\iota_{m,n}$ pro $n \geq m$. □

Příklad 3.11. Pseudovarieta $\mathbf{N}$ je zadaná posloupností identit $x_1 \ldots x_k = 0$ a pseudovarietu $\mathbf{G}$ můžeme zadat posloupností identit $x^{k!} = 1$.

3.2 Pseudoidentity

Jak jsme viděli v příkladu 3.5 a ve cvičení 3.7, přestože pseudovariety není možné obecně popsat pomocí obvyklých identit, je možné v některých případech k jejich popisu použít zobecněné identity používající operaci umocnění na ω. Nyní popíšeme, jaké zobecnění identit musíme uvažovat, aby bylo možné libovolnou pseudovarietu popsat pomocí systému rovností. Těmto zobecněným identitám se říká pseudoidentity, přičemž na každé straně pseudoidentity stojí takzvaná implicitní operace. Jedna implicitní operace dané arity sestává z jedné operace této arity pro každý konečný monoid, přičemž tyto jednotlivé operace musejí respektovat homomorfismy.

Definice 3.12. Nechť $\mathbf{M}$ značí pseudovarietu všech konečných monoidů. Buď $n \in \mathbb{N}_0$. Systém zobrazení $\pi = (\pi_S)_{S\in\mathbf{M}}$, kde $\pi_S \colon S^n \to S$, se nazývá *n-ární implicitní operace*, jestliže pro všechny konečné monoidy S a T, pro všechny homomorfismy $\varphi \colon S \to T$ a pro všechny prvky $x_1,\ldots,x_n \in S$ platí

$$\varphi(\pi_S(x_1,\ldots,x_n)) = \pi_T(\varphi(x_1),\ldots,\varphi(x_n)),$$

tedy jestliže komutuje následující diagram:

$$\begin{tikzcd} S^n \arrow[r, "\pi_S"] \arrow[d, "\varphi^n"'] & S \arrow[d, "\varphi"] \\ T^n \arrow[r, "\pi_T"] & T \end{tikzcd}$$

Množinu všech n-árních implicitních operací budeme značit $\mathrm{I}(n)$.

Všimněte si, že podmínka na kompatibilitu s homomorfismy mimo jiné zaručuje, že implicitní operaci stačí zadat na jednom monoidu z každé třídy vzájemně izomorfních

monoidů, tedy dohromady na spočetně mnoha monoidech. Z toho vyplývá, že pro každé n má množina $\mathrm{I}(n)$ nejvýše mohutnost kontinua. Dá se dokázat, že dokonce již množina $\mathrm{I}(1)$ skutečně mohutnost kontinua má.

Je-li $A = \{a_1, \dots, a_n\}$ n-prvková abeceda proměnných, zadává každé slovo $u \in A^*$ na libovolném konečném monoidu S n-ární operaci u_S tak, že hodnotu $u_S(x_1, \dots, x_n)$ pro $x_1, \dots, x_n \in S$ získáme dosazením každého z prvků x_i za příslušnou proměnnou a_i. Z univerzální algebry víme, že potom $(u_S)_{S \in \mathbf{M}}$ je implicitní operace, které se říká *termová operace* (uvědomte si, že slova jsou vlastně termy s jedním binárním operačním symbolem, uvažované modulo asociativita, kterou v celé teorii implicitně předpokládáme). Přitom dvě různá slova $u, v \in A^*$ nemohou reprezentovat tutéž implicitní operaci, neboť existuje konečný monoid S (a podle příkladu 3.5 dokonce grupa), který nesplňuje identitu $u = v$, a tedy se na něm operace u_S a v_S liší. Proto můžeme podmnožinu všech termových operací v $\mathrm{I}(n)$ ztotožnit s množinou A^*. Všimněte si, že každé písmeno $a_i \in A$, chápané jako implicitní operace, představuje právě projekci na i-tou složku.

Cvičení 3.13. Dokažte, že předpis $\pi_S(x) = x^\omega$ pro $x \in S$ definuje unární implicitní operaci.

Cvičení 3.14. Dokažte, že jsou-li $\pi_1, \dots, \pi_n$ m-ární implicitní operace a π je n-ární implicitní operace, potom kompozice $\pi(\pi_1, \dots, \pi_n)$ definovaná předpisem

$$(\pi(\pi_1, \dots, \pi_n))_S(x_1, \dots, x_m) = \pi_S((\pi_1)_S(x_1, \dots, x_m), \dots, (\pi_n)_S(x_1, \dots, x_m))$$

je m-ární implicitní operace. Podobně jako termové operace tedy i implicitní operace na každém monoidu zadávají klon funkcí.

Mimo jiné tedy máme pro libovolné n-ární implicitní operace π a σ definovánu n-ární implicitní operaci $\pi \cdot \sigma$ předpisem

$$(\pi \cdot \sigma)_S(x_1, \dots, x_n) = \pi_S(x_1, \dots, x_n) \cdot \sigma_S(x_1, \dots, x_n).$$

Cvičení 3.15. Dokažte, že vzhledem k takto definované operaci násobení tvoří $\mathrm{I}(n)$ monoid.

Na monoidu $\mathrm{I}(n)$ nyní definujeme metriku, která vyjadřuje, jak velký konečný monoid potřebujeme k tomu, abychom dané implicitní operace odlišili. Pro libovolné různé implicitní operace $\pi, \sigma \in \mathrm{I}(n)$ tedy označme r počet prvků nejmenšího monoidu S takového, že $\pi_S \neq \sigma_S$. Protože čím větší monoid k rozlišení potřebujeme, tím jsou si implicitní operace podobnější, položíme $d(\pi, \sigma) = 2^{-r}$.

Cvičení 3.16. Dokažte, že d je ultrametrika na $\mathrm{I}(n)$, tedy metrika splňující silnější variantu trojúhelníkové nerovnosti:

$$\forall \pi, \sigma, \tau \in \mathrm{I}(n) \colon d(\pi, \tau) \leq \max(d(\pi, \sigma), d(\sigma, \tau)).$$

Cvičení 3.17. Ukažte, že posloupnost n-árních implicitních operací $(\pi_k)_{k=1}^{\infty}$ konverguje k implicitní operaci π právě tehdy, když pro každý konečný monoid S existuje $n \in \mathbb{N}$ takové, že pro všechna $k \geq n$ platí $(\pi_k)_S = \pi_S$.

Příklad 3.18. Posloupnost unárních termových operací $a_1^{k!}$ konverguje pro $k \to \infty$ k implicitní operaci a_1^{ω}.

Cvičení 3.19. Dokažte, že operace násobení implicitních operací je spojité zobrazení $\mathrm{I}(n) \times \mathrm{I}(n) \to \mathrm{I}(n)$. Můžeme tedy říct, že $\mathrm{I}(n)$ je metrický monoid.

Cvičení 3.20. Dokažte, že metrický prostor $\mathrm{I}(n)$ je úplný.

Tvrzení 3.21. *Metrický prostor* $\mathrm{I}(n)$ *je kompaktní.*

Důkaz. Nechť S_k, pro $k \in \mathbb{N}$, jsou až na izomorfismus všechny konečné monoidy. Buď $(\pi_k)_{k=1}^{\infty}$ libovolná posloupnost n-árních implicitních operací. Ukážeme, že tato posloupnost obsahuje podposloupnost $(\pi_{i_k})_{k=1}^{\infty}$ takovou, že pro všechna $k \in \mathbb{N}$ se implicitní operace π_{i_ℓ} pro $\ell \geq k$ shodují na S_k. Tato podposloupnost je cauchyovská, a proto je podle cvičení 3.20 konvergentní, což dokazuje kompaktnost $\mathrm{I}(n)$.

Hledanou podposloupnost zkonstruujme induktivně. Předpokládejme, že po m-tém kroku konstrukce máme podposloupnost

$$\pi_{i_1}, \ldots, \pi_{i_m}, \pi_{i_{m,1}}, \pi_{i_{m,2}}, \ldots$$

takovou, že pro všechna $k \in \{1, \ldots, m\}$ se implicitní operace π_{i_ℓ} pro $\ell \in \{k, \ldots, m\}$ a $\pi_{i_{m,\ell}}$ pro $\ell \in \mathbb{N}$ shodují na S_k. Tedy známe prvních m členů hledané podposloupnosti a ze zbytku posloupnosti máme vybranou vhodnou podposloupnost tak, aby byla splněna podmínka, kterou po hledané podposloupnosti požadujeme. Protože existuje pouze konečně mnoho různých n-árních operací na monoidu S_{m+1}, nekonečně mnoho prvků posloupnosti $(\pi_{i_{m,\ell}})_{\ell=1}^{\infty}$ se na tomto monoidu shoduje. Zvolíme-li nyní

$$\pi_{i_{m+1}}, \pi_{i_{m+1,1}}, \pi_{i_{m+1,2}}, \ldots$$

jako podposloupnost posloupnosti $(\pi_{i_{m,\ell}})_{\ell=1}^{\infty}$ složenou z prvků shodujících se na S_{m+1}, získáme posloupnost splňující předpoklady pro další krok konstrukce. □

Poznámka 3.22. Je-li $\mathscr{M}$ spočetná množina obsahující jeden monoid z každé třídy vzájemně izomorfních konečných monoidů, potom můžeme množinu implicitních operací $\mathrm{I}(n)$ chápat jako podmnožinu množiny $\prod_{S \in \mathscr{M}} S^{S^n}$. Uvažujeme-li na každém monoidu $S \in \mathscr{M}$ diskrétní topologii, indukuje námi definovaná metrika na $\mathrm{I}(n)$ právě topologii podprostoru prostoru $\prod_{S \in \mathscr{M}} S^{S^n}$ vybaveného součinovou topologií, který je izomorfní Cantorovu diskontinuu. Kompaktnost $\mathrm{I}(n)$ tedy vyplývá z toho, že je úplný, a tedy uzavřený, podprostor kompaktního prostoru.

Cvičení 3.23. Ukažte, že A^* je diskrétní podprostor $\mathrm{I}(n)$, tedy že pro každou termovou operaci $\pi = (u_S)_{S\in\mathbf{M}}$ existuje $\varepsilon > 0$ takové, že žádná termová operace nemá vzdálenost od π menší než ε.

Tvrzení 3.24. *Každá implicitní operace $\pi \in \mathrm{I}(n)$ je limitou konvergentní posloupnosti termových operací.*

Důkaz. Potřebujeme dokázat, že pro libovolné $r \in \mathbb{N}$ existuje nějaké slovo $u \in A^*$ takové, že $\pi_S = u_S$ pro všechny monoidy S mohutnosti menší než r. Za tímto účelem ukážeme, že implicitní operace π je na všech těchto monoidech určena obrazem jediné n-tice na nějakém mnohem větším monoidu. Nechť $\mathscr{M}$ je (konečná) množina obsahující právě jeden monoid z každé třídy vzájemně izomorfních monoidů o méně než r prvcích. Uvažujme podmonoid T generovaný v (konečném) monoidu $\prod_{S\in\mathscr{M}} S^{S^n}$ prvky, které jsou na každé složce rovny i-té projekci, tedy prvky

$$y_i = (x_i)_{S\in\mathscr{M},\ (x_1,\dots,x_n)\in S^n}$$

pro $i = 1,\dots,n$. Protože T je generovaný prvky $y_1, \dots, y_n$, existuje nějaké slovo $u \in A^*$ takové, že $\pi_T(y_1,\dots,y_n) = u_T(y_1,\dots,y_n)$. Myšlenka právě provedené konstrukce spočívá v tom, že výsledek operace π_T na n-tici $y_1, \dots, y_n$ v sobě kóduje výsledky operací π_S na všech n-ticích prvků monoidu S pro všechny monoidy $S \in \mathscr{M}$, přičemž hodnota $\pi_S(x_1,\dots,x_n)$ je uvedena na složce prvku $\pi_T(y_1,\dots,y_n)$ s indexy S, $(x_1,\dots,x_n)$.

Zbývá ukázat, že pro libovolný monoid $R \in \mathscr{M}$ platí $\pi_R = u_R$. Za tímto účelem vezměme libovolnou n-tici $(z_1,\dots,z_n) \in R^n$. Tato n-tice je obrazem n-tice $(y_1,\dots,y_n)$ v projekci $\mathsf{p}_{R,\ (z_1,\dots,z_n)}\colon T \to R$, kterou označíme p. Jelikož každá projekce je homomorfismus, můžeme pomocí vlastností implicitní operace π spočítat:

$$\begin{aligned}\pi_R(z_1,\dots,z_n) &= \pi_R(p(y_1),\dots,p(y_n)) = p(\pi_T(y_1,\dots,y_n)) = p(u_T(y_1,\dots,y_n))\\ &= u_R(p(y_1),\dots,p(y_n)) = u_R(z_1,\dots,z_n).\end{aligned}$$

□

Všimněte si, že cvičení 3.20 a tvrzení 3.24 dohromady ukazují, že prostor $\mathrm{I}(n)$ je zúplněním metrického prostoru A^*. Tento fakt umožňuje metrický monoid $\mathrm{I}(n)$ ekvivalentně definovat tak, že metriku d zavedeme pouze na A^* a vzniklý metrický monoid zúplníme.

Chápejme odteď každý konečný monoid S jako diskrétní metrický monoid, tedy například s metrikou $d(x,y) = 1$ pro všechna $x,y \in S$, $x \neq y$. Tvrzení 3.24 zaručuje, že každý spojitý homomorfismus $\varphi\colon \mathrm{I}(n) \to S$, kde S je libovolný metrický monoid, je jednoznačně určen obrazy písmen abecedy A: na podmonoid A^* se jednoznačně rozšíří, neboť se jedná o homomorfismus, a na celý monoid $\mathrm{I}(n)$ se rozšíří díky hustotě A^* v $\mathrm{I}(n)$. Pro konečný monoid S dokonce existuje vzájemně jednoznačná korespondence mezi zobrazeními $A \to S$ a spojitými homomorfismy $\mathrm{I}(n) \to S$, přičemž každý z těchto homomorfismů funguje jako dosazování obrazů písmen do implicitních operací:

Tvrzení 3.25. *Je-li S konečný monoid a $\varphi\colon \mathrm{I}(n) \to S$ spojitý homomorfismus, potom pro všechny implicitní operace $\pi \in \mathrm{I}(n)$ platí*

$$\varphi(\pi) = \pi_S(\varphi(a_1), \dots, \varphi(a_n)).$$

Naopak, pro každé zobrazení $f\colon A \to S$ zadává předpis

$$\varphi(\pi) = \pi_S(f(a_1), \dots, f(a_n))$$

spojitý homomorfismus $\varphi\colon \mathrm{I}(n) \to S$ splňující $\varphi(a_i) = f(a_i)$ pro $i = 1, \dots, n$.

Cvičení 3.26. Dokažte předchozí tvrzení.

Spojité homomorfismy z metrického monoidu $\mathrm{I}(n)$ ve skutečnosti hrají v teorii pseudovariet stejnou úlohu jako homomorfismy z monoidu A^* v teorii variet.

Pseudoidentitou rozumíme libovolnou dvojici implicitních operací stejné arity $(\pi, \sigma) \in \mathrm{I}(n) \times \mathrm{I}(n)$, obvykle zapisovanou jako $\pi = \sigma$. Říkáme, že konečný monoid S *splňuje* pseudoidentitu $\pi = \sigma$ a píšeme $S \models \pi = \sigma$, jestliže $\pi_S = \sigma_S$.

Lemma 3.27. *Konečný monoid S splňuje pseudoidentitu $\pi = \sigma$, kde $\pi, \sigma \in \mathrm{I}(n)$, právě tehdy, když pro všechny spojité homomorfismy $\varphi\colon \mathrm{I}(n) \to S$ platí $\varphi(\pi) = \varphi(\sigma)$.*

Důkaz. Stačí si všimnout, že podle tvrzení 3.25 je požadavek, aby platilo $\pi_S(x_1, \dots, x_n) = \sigma_S(x_1, \dots, x_n)$ pro všechny prvky $x_1, \dots, x_n \in S$, ekvivalentní požadavku

$$\pi_S(\varphi(a_1), \dots, \varphi(a_n)) = \sigma_S(\varphi(a_1), \dots, \varphi(a_n)),$$

neboli $\varphi(\pi) = \varphi(\sigma)$, pro všechny spojité homomorfismy $\varphi\colon \mathrm{I}(n) \to S$. □

Pro množinu pseudoidentit Π říkáme, že S splňuje Π a píšeme $S \models \Pi$, jestliže splňuje všechny pseudoidentity z Π. Třídu všech konečných monoidů splňujících Π značíme $\mathrm{Mod}(\Pi)$.

Nechť $\mathbf{V}$ je pseudovarieta konečných monoidů. Kongruenci $\sim_{\mathbf{V}}^{A}$ na monoidu A^* rozšíříme na monoid $\mathrm{I}(n)$ tak, že položíme $\pi \sim_{\mathbf{V}}^{n} \sigma$, jestliže $S \models \pi = \sigma$ platí pro všechny monoidy $S \in \mathbf{V}$. Tedy $\sim_{\mathbf{V}}^{n}$ sestává ze všech n-árních pseudoidentit splněných ve $\mathbf{V}$.

Cvičení 3.28. Dokažte, že $\sim_{\mathbf{V}}^{n}$ je kongruence monoidu $\mathrm{I}(n)$, která je uzavřenou podmnožinou metrického prostoru $\mathrm{I}(n) \times \mathrm{I}(n)$.

Na množině $\mathrm{I}(n)/{\sim_{\mathbf{V}}^{n}}$ definujeme metriku $d_{\mathbf{V}}$ analogicky metrice na $\mathrm{I}(n)$. Pro $\pi, \sigma \in \mathrm{I}(n)$ tedy označme r počet prvků nejmenšího monoidu $S \in \mathbf{V}$, který nesplňuje pseudoidentitu $\pi = \sigma$, a definujme $d_{\mathbf{V}}(\pi{\sim_{\mathbf{V}}^{n}}, \sigma{\sim_{\mathbf{V}}^{n}}) = 2^{-r}$.

Přirozená projekce $\nu\colon \mathrm{I}(n) \to \mathrm{I}(n)/{\sim_{\mathbf{V}}^{n}}$ je podle cvičení 3.28 homomorfismus a protože pro všechny $\pi, \sigma \in \mathrm{I}(n)$ platí $d_{\mathbf{V}}(\pi{\sim_{\mathbf{V}}^{n}}, \sigma{\sim_{\mathbf{V}}^{n}}) \leq d(\pi, \sigma)$, je navíc spojitá. Jako spojitý obraz kompaktního prostoru je tedy $\mathrm{I}(n)/{\sim_{\mathbf{V}}^{n}}$ kompaktní.

V následujících důkazech využijeme jednoduché tvrzení o spojitých zobrazeních z kompaktního prostoru:

Cvičení 3.29. Nechť M, N a P jsou metrické prostory, přičemž M je kompaktní. Nechť dále $\varphi\colon M \to N$ a $\rho\colon M \twoheadrightarrow P$ jsou spojitá zobrazení, přičemž ρ je surjektivní. Dokažte, že platí-li $\ker(\rho) \subseteq \ker(\varphi)$, tak jedinečné zobrazení $\psi\colon P \to N$ splňující $\psi \circ \rho = \varphi$ je spojité.

Toto tvrzení nám umožňuje převést testování, zda monoid $S \in \mathbf{V}$ splňuje pseudoidentitu $\pi = \sigma$, pomocí spojitých homomorfismů z monoidu $\mathrm{I}(n)$ (tj. lemma 3.27) na testování pomocí spojitých homomorfismů z monoidu $\mathrm{I}(n)/{\sim}^n_{\mathbf{V}}$:

Lemma 3.30. *Konečný monoid S náležející do pseudovariety $\mathbf{V}$ splňuje pseudoidentitu $\pi = \sigma$, kde $\pi, \sigma \in \mathrm{I}(n)$, právě tehdy, když pro všechny spojité homomorfismy $\psi\colon \mathrm{I}(n)/{\sim}^n_{\mathbf{V}} \to S$ platí $\psi(\pi{\sim}^n_{\mathbf{V}}) = \psi(\sigma{\sim}^n_{\mathbf{V}})$.*

Důkaz. Předpokládejme, že monoid $S \in \mathbf{V}$ splňuje pseudoidentitu $\pi = \sigma$ a $\psi\colon \mathrm{I}(n)/{\sim}^n_{\mathbf{V}} \to S$ je libovolný spojitý homomorfismus. Složením s přirozenou projekcí $\nu\colon \mathrm{I}(n) \to \mathrm{I}(n)/{\sim}^n_{\mathbf{V}}$ dostaneme spojitý homomorfismus $\psi \circ \nu\colon \mathrm{I}(n) \to S$, takže podle lemmatu 3.27 platí rovnost $\psi(\nu(\pi)) = \psi(\nu(\sigma))$, kterou jsme chtěli ověřit.

Opačně, předpokládejme, že rovnost $\psi(\pi{\sim}^n_{\mathbf{V}}) = \psi(\sigma{\sim}^n_{\mathbf{V}})$ platí pro všechny spojité homomorfismy $\psi\colon \mathrm{I}(n)/{\sim}^n_{\mathbf{V}} \to S$. K ověření, že monoid S splňuje pseudoidentitu $\pi = \sigma$, použijeme lemma 3.27. Buď tedy $\varphi\colon \mathrm{I}(n) \to S$ libovolný spojitý homomorfismus. Potom platí $\ker(\nu) \subseteq \ker(\varphi)$, jelikož pro libovolné *kappa*, $\lambda \in \mathrm{I}(n)$ z rovnosti $\nu(\kappa) = \nu(\lambda)$, tj. $\kappa \sim^n_{\mathbf{V}} \lambda$, díky předpokladu, že monoid S patří do $\mathbf{V}$, plyne $S \models \kappa = \lambda$, takže podle lemmatu 3.27 je splněno $\varphi(\kappa) = \varphi(\lambda)$. Cvičení 3.29 nám tedy garantuje existenci spojitého homomorfismu $\psi\colon \mathrm{I}(n)/{\sim}^n_{\mathbf{V}} \to S$ takového, že $\psi \circ \nu = \varphi$. Z předpokládané rovnosti $\psi(\pi{\sim}^n_{\mathbf{V}}) = \psi(\sigma{\sim}^n_{\mathbf{V}})$ proto dostáváme $\varphi(\pi) = \psi(\nu(\pi)) = \psi(\nu(\sigma)) = \varphi(\sigma)$. Podle lemmatu 3.27 tedy monoid S splňuje pseudoidentitu $\pi = \sigma$. □

Lemma 3.31. *Pokud konečný monoid S splňuje všechny pseudoidentity platné v pseudovarietě $\mathbf{V}$, tak je spojitým homomorfním obrazem monoidu $\mathrm{I}(n)/{\sim}^n_{\mathbf{V}}$ pro nějaké $n \in \mathbb{N}$.*

Důkaz. Podle tvrzení 3.25 existuje surjektivní spojitý homomorfismus $\varphi\colon \mathrm{I}(n) \twoheadrightarrow S$ pro dostatečně velké n (například $n = |S|$). Jelikož S splňuje všechny pseudoidentity patřící do ${\sim}^n_{\mathbf{V}}$, z lemmatu 3.27 dostáváme ${\sim}^n_{\mathbf{V}} \subseteq \ker(\varphi)$. Podle cvičení 3.29 tedy předpis $\psi(\pi{\sim}^n_{\mathbf{V}}) = \varphi(\pi)$ zadává spojitý homomorfismus $\psi\colon \mathrm{I}(n)/{\sim}^n_{\mathbf{V}} \twoheadrightarrow S$. □

Tvrzení 3.32. *Pro každou pseudovarietu $\mathbf{V}$ platí, že konečný monoid patří do $\mathbf{V}$ právě tehdy, když je spojitým homomorfním obrazem monoidu $\mathrm{I}(n)/{\sim}^n_{\mathbf{V}}$ pro nějaké $n \in \mathbb{N}$.*

Důkaz. Přímá implikace vyplývá triviálně z předchozího lemmatu.

Předpokládejme tedy, že $\varphi\colon \mathrm{I}(n)/{\sim}^n_{\mathbf{V}} \twoheadrightarrow S$ je surjektivní spojitý homomorfismus. Protože prostor $\mathrm{I}(n)/{\sim}^n_{\mathbf{V}}$ je kompaktní, je zobrazení φ stejnoměrně spojité. Díky konečnosti monoidu S tedy existuje $\delta > 0$ takové, že pro všechny $\pi, \sigma \in \mathrm{I}(n)$ splňující $d_{\mathbf{V}}(\pi{\sim}^n_{\mathbf{V}}, \sigma{\sim}^n_{\mathbf{V}}) < \delta$ platí $\varphi(\pi{\sim}^n_{\mathbf{V}}) = \varphi(\sigma{\sim}^n_{\mathbf{V}})$. Zvolme $r \in \mathbb{N}$ takové, aby splňovalo

$2^{-r} \leq \delta$. Nechť $\psi_k\colon \mathrm{I}(n)/{\sim_{\mathbf{V}}^n} \to T_k$, pro $k = 1,\dots,m$, jsou až na izomorfismus všechny spojité homomorfismy do monoidů patřících do $\mathbf{V}$, které mají nejvýše r prvků; těchto homomorfismů je skutečně jen konečně mnoho, jelikož každý takový spojitý homomorfismus je jednoznačně zadaný obrazy písmen. Uvažujme spojitý homomorfismus $\psi\colon \mathrm{I}(n)/{\sim_{\mathbf{V}}^n} \to \prod_{k=1}^m T_k$ definovaný předpisem $(\psi(\pi{\sim_{\mathbf{V}}^n}))_k = \psi_k(\pi{\sim_{\mathbf{V}}^n})$.

Ukažme nyní, že platí $\ker(\psi) \subseteq \ker(\varphi)$. Pokud nějaké $\pi,\sigma \in \mathrm{I}(n)$ splňují rovnost $\psi(\pi{\sim_{\mathbf{V}}^n}) = \psi(\sigma{\sim_{\mathbf{V}}^n})$, znamená to, že $\psi_k(\pi{\sim_{\mathbf{V}}^n}) = \psi_k(\sigma{\sim_{\mathbf{V}}^n})$ pro $k = 1,\dots,m$. Podle lemmatu 3.30 tedy všechny monoidy ve $\mathbf{V}$ o nejvýše r prvcích splňují pseudoidentitu $\pi = \sigma$. Proto platí $d_{\mathbf{V}}(\pi{\sim_{\mathbf{V}}^n}, \sigma{\sim_{\mathbf{V}}^n}) < 2^{-r} \leq \delta$, což díky volbě δ dává požadovanou rovnost $\varphi(\pi{\sim_{\mathbf{V}}^n}) = \varphi(\sigma{\sim_{\mathbf{V}}^n})$.

Díky surjektivitě homomorfismu φ z platnosti $\ker(\psi) \subseteq \ker(\varphi)$ plyne, že monoid S je homomorfním obrazem podmonoidu $\psi(\mathrm{I}(n)/{\sim_{\mathbf{V}}^n})$ součinu monoidů T_k patřících do $\mathbf{V}$, a tedy sám patří do $\mathbf{V}$. □

Věta 3.33 (Jan Reiterman, 1982)**.** *Třída konečných monoidů* $\mathbf{V}$ *je pseudovarieta právě tehdy, když existuje množina pseudoidentit* Π *taková, že* $\mathbf{V} = \mathrm{Mod}(\Pi)$.

Důkaz. Snadno se ověří, podobně jako v případě variet, že $\mathrm{Mod}(\Pi)$ je vždy pseudovarieta.

Opačně, je-li $\mathbf{V}$ pseudovarieta, zvolme za Π množinu všech pseudoidentit platných ve $\mathbf{V}$, tedy $\bigcup_{n\in\mathbb{N}} {\sim_{\mathbf{V}}^n}$. Potom zřejmě platí $\mathbf{V} \subseteq \mathrm{Mod}(\Pi)$. Naopak, libovolný konečný monoid splňující všechny pseudoidentity platné ve $\mathbf{V}$ je podle lemmatu 3.31 spojitým homomorfním obrazem $\mathrm{I}(n)/{\sim_{\mathbf{V}}^n}$, a tedy podle tvrzení 3.32 patří do $\mathbf{V}$. □

Podobně je možné dokázat, že pseudovariety konečných uspořádaných monoidů jsou právě třídy definovatelné pomocí pseudoidentit tvaru $\pi \leq \sigma$ pro $\pi,\sigma \in \mathrm{I}(n)$.

Použití pseudoidentit si ukážeme na pseudovarietách všech konečných $\mathscr{R}$-triviálních monoidů $\mathbf{R}$ a všech konečných $\mathscr{J}$-triviálních monoidů $\mathbf{J}$.

Cvičení 3.34. Dokažte, že $\mathbf{R} = \mathrm{Mod}((xy)^\omega x = (xy)^\omega)$.

Protože konečný monoid je $\mathscr{J}$-triviální právě tehdy, když je současně $\mathscr{R}$-triviální a $\mathscr{L}$-triviální, dostáváme z předchozího cvičení, že

$$\mathbf{J} = \mathrm{Mod}((xy)^\omega x = (xy)^\omega = y(xy)^\omega). \tag{3.1}$$

Příklad 3.35. Ukážeme, že platí

$$\mathbf{J} = \mathrm{Mod}(x^{\omega+1} = x^\omega,\ (xy)^\omega = (yx)^\omega). \tag{3.2}$$

Všimněte si, že pseudoidentita $x^{\omega+1} = x^\omega$ popisuje aperiodické (ekvivalentně $\mathscr{H}$-triviální) monoidy; tato pseudoidentita plyne z pseudoidentity $(xy)^\omega x = (xy)^\omega$ dosazením 1 za y. Druhou pseudoidentitu potom získáme následovně:

$$(xy)^\omega = (xy)^\omega x = x(yx)^\omega = (yx)^\omega.$$

Opačně, z platnosti pseudoidentit $x^{\omega+1} = x^{\omega}$ a $(xy)^{\omega} = (yx)^{\omega}$ nejprve odvodíme

$$(xy)^{\omega} = (yx)^{\omega} = (yx)^{\omega+1} = y(xy)^{\omega}x.$$

Indukcí z tohoto vztahu dostaneme $(xy)^{\omega} = y^{\omega}(xy)^{\omega}x^{\omega}$, což nám umožní použít pseudoidentitu $x^{\omega+1} = x^{\omega}$ ke spočítaní

$$(xy)^{\omega} = y^{\omega}(xy)^{\omega}x^{\omega} = y^{\omega}(xy)^{\omega}x^{\omega} \cdot x = (xy)^{\omega}x.$$

Druhá pseudoidentita v (3.1) se odvodí symetricky.

3.3 Eilenbergova korespondence

V této kapitole si ukážeme, že třídy regulárních jazyků, které je možné popsat pomocí pseudovariet monoidů, jsou právě takzvané variety. Nejprve si ale musíme ujasnit, co je to vlastně třída regulárních jazyků. Nemůžeme ji totiž chápat jako třídu množin slov ve smyslu teorie množin, neboť v případě mnoha přirozených tříd jazyků závisí příslušnost jazyka do této třídy nejen na jazyku samotném, ale i na tom, nad jakou abecedou jej právě uvažujeme. Proto chápeme třídu jazyků jako zobrazení, které každé konečné abecedě přiřazuje nějakou množinu jazyků nad touto abecedou. Přitom můžeme předpokládat, že konečných abeced je jen množina, neboť se nebudeme zabývat třídami, u nichž se příslušnost jazyka může změnit přejmenováním písmen, a tedy si vystačíme s jedinou množinou písmen pro každou velikost abecedy.

Definice 3.36. *Varieta regulárních jazyků* je zobrazení $\mathscr{L}$, které každé konečné abecedě A přiřazuje nějakou neprázdnou množinu regulárních jazyků $\mathscr{L}(A) \subseteq \mathscr{P}(A^*)$ a přitom pro všechny konečné abecedy A a B splňuje následující podmínky:

1. Množina $\mathscr{L}(A)$ je uzavřená na booleovské operace, tedy na konečná sjednocení a průniky a na komplementy v A^*.
2. Pokud $L \in \mathscr{L}(A)$, potom pro všechna $a \in A$ platí $a^{-1} \cdot L \in \mathscr{L}(A)$ a $L \cdot a^{-1} \in \mathscr{L}(A)$.
3. Je-li $\psi\colon B^* \to A^*$ homomorfismus a platí $L \in \mathscr{L}(A)$, potom $\psi^{-1}(L) \in \mathscr{L}(B)$.

Cvičení 3.37. Dokažte, že je-li nějaká třída regulárních jazyků uzavřená na kvocienty písmeny a na konečná sjednocení, je uzavřená i na kvocienty libovolným jazykem.

Variety jazyků jsou přirozeně uspořádány inkluzí po složkách, tedy $\mathscr{K} \leq \mathscr{L}$, jestliže pro všechny konečné abecedy A platí $\mathscr{K}(A) \subseteq \mathscr{L}(A)$.

Věta 3.38 (Samuel Eilenberg, 1976)**.** *Nechť pro každou pseudovarietu konečných monoidů* $\mathbf{V}$ *značí* $\lambda(\mathbf{V})$ *varietu regulárních jazyků definovanou předpisem*

$$\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid M_L \in \mathbf{V}\}.$$

Nechť dále pro každou varietu regulárních jazyků $\mathscr{L}$ značí $\mu(\mathscr{L})$ pseudovarietu konečných monoidů generovanou všemi syntaktickými monoidy M_L, kde L je regulární jazyk nad nějakou konečnou abecedou A splňující $L \in \mathscr{L}(A)$. Potom λ a μ jsou vzájemně inverzní izomorfismy mezi svazem všech pseudovariet konečných monoidů a svazem všech variet regulárních jazyků.

Cvičení 3.39. Ukažte, že $\lambda(\mathbf{V})$ je skutečně varieta regulárních jazyků. Uvědomte si, že zobrazení λ je možné ekvivalentně definovat předpisem

$$\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid L \text{ je rozpoznáván nějakým homomorfismem } \varphi\colon A^* \to S, \text{ kde } S \in \mathbf{V}\}.$$

Důkaz. Jelikož zobrazení λ a μ jsou zřejmě izotonní, stačí ukázat, že se jedná o vzájemně inverzní bijekce.

Pro libovolnou pseudovarietu $\mathbf{V}$ je inkluze $\mu\lambda(\mathbf{V}) \subseteq \mathbf{V}$ zřejmá. Abychom ukázali opačnou inkluzi, uvažme libovolný monoid $S \in \mathbf{V}$. Jelikož je tento monoid konečný, existuje surjektivní homomorfismus $\varphi\colon A^* \twoheadrightarrow S$ pro nějakou konečnou abecedu A. Pro každý prvek $x \in S$ označme M_x syntaktický monoid jazyka $\varphi^{-1}(x) \in \lambda(\mathbf{V})(A)$. Příslušnost monoidu S do pseudovariety $\mu\lambda(\mathbf{V})$ nyní ověříme tak, že ukážeme, že je izomorfní podmonoidu součinu monoidů M_x, pro $x \in S$. Z vlastností syntaktického monoidu víme (viz diagram (1.1)), že existuje homomorfismus $\psi_x\colon S \to M_x$ takový, že $\psi_x \circ \varphi$ je syntaktický homomorfismus jazyka $\varphi^{-1}(x)$. Platí tedy $\varphi^{-1}(\psi_x^{-1}(\psi_x(x))) = \varphi^{-1}(x)$, což díky surjektivitě φ znamená, že $\psi_x^{-1}(\psi_x(x)) = \{x\}$. Homomorfismus $\sigma\colon S \to \prod_{x\in S} M_x$ definujeme předpisem $\sigma(y) = (\psi_x(y))_{x\in S}$. Z předchozího pozorování plyne, že tento homomorfismus je skutečně injektivní, neboť pro libovolné $y, z \in S$, $y \neq z$, platí $\psi_y(y) \neq \psi_y(z)$, a tedy se $\sigma(y)$ liší od $\sigma(z)$ na složce odpovídající y. Proto platí $S \in \mu\lambda(\mathbf{V})$.

Zbývá ukázat, že pro libovolnou varietu $\mathscr{L}$ platí $\lambda\mu(\mathscr{L}) = \mathscr{L}$. Přitom pro libovolnou konečnou abecedu A je inkluze $\mathscr{L}(A) \subseteq (\lambda\mu(\mathscr{L}))(A)$ triviální. Abychom ověřili opačnou inkluzi, dokážeme, že všechny jazyky nad abecedou A rozpoznávané nějakým monoidem z pseudovariety $\mu(\mathscr{L})$ patří do $\mathscr{L}(A)$. Každý monoid z $\mu(\mathscr{L})$ lze ovšem vytvořit z monoidů M_L, kde $L \in \mathscr{L}(B)$, konstruováním homomorfních obrazů, podmonoidů a konečných součinů. Přitom jazyky rozpoznávané homomorfním obrazem či podmonoidem daného monoidu jsou vždy rozpoznávané i původním monoidem. Dále, každý jazyk rozpoznávaný součinem $S \times T$ je konečným sjednocením jazyků tvaru $\varphi^{-1}((x,y)) = (p_1 \circ \varphi)^{-1}(x) \cap (p_2 \circ \varphi)^{-1}(y)$ a lze tedy získat z jazyků rozpoznávaných S nebo T pomocí operací, na které je množina $\mathscr{L}(A)$ uzavřená. K dokončení důkazu nám tedy stačí ověřit, že všechny jazyky nad A rozpoznávané syntaktickým monoidem M_L libovolného jazyka $L \in \mathscr{L}(B)$ nad libovolnou abecedou B patří do $\mathscr{L}(A)$.

Uvažujme tedy libovolný homomorfismus $\varphi\colon A^* \to M_L$ a podmnožinu $F \subseteq M_L$. Jelikož syntaktický homomorfismus φ_L je surjektivní a monoid A^* je volný, existuje nějaký homomorfismus $\psi\colon A^* \to B^*$ splňující $\varphi_L \circ \psi = \varphi$. Potom ovšem platí $\varphi^{-1}(F) = \psi^{-1}(\varphi_L^{-1}(F))$, a proto díky uzavřenosti $\mathscr{L}$ na vzory v homomorfismech stačí k dokázání $\varphi^{-1}(F) \in \mathscr{L}(A)$ ověřit $\varphi_L^{-1}(F) \in \mathscr{L}(B)$. K tomu je dostačující ukázat,

že pro každý prvek $x \in M_L$ patří $\varphi_L^{-1}(x)$ do $\mathscr{L}(B)$. Přitom slova jazyka $\varphi_L^{-1}(x)$ jsou jednoznačně popsána množinou svých kontextů v jazyce L. Je-li tedy w libovolné slovo z $\varphi_L^{-1}(x)$, platí

$$\varphi_L^{-1}(x) = \bigcap_{(u,v)\in \mathrm{C}_L(w)} u^{-1}\cdot L\cdot v^{-1} \cap \bigcap_{(u,v)\notin \mathrm{C}_L(w)} \overline{u^{-1}\cdot L\cdot v^{-1}}.$$

Jelikož podle věty 1.21 existuje pro regulární jazyk L jen konečně mnoho jazyků tvaru $u^{-1}\cdot L\cdot v^{-1}$, jsou oba průniky v tomto vyjádření ve skutečnosti konečné, takže jazyk $\varphi_L^{-1}(x)$ lze opravdu z jazyka L získat pomocí operací, na které je množina $\mathscr{L}(B)$ uzavřená. □

Analogickou větu je možné dokázat i pro pseudovariety konečných pologrup, které odpovídají varietám regulárních jazyků neobsahujících prázdné slovo. Definice těchto variet se od definice 3.36 liší pouze tím, že uzavřenost na vzory požadujeme pouze pro homomorfismy, které nezobrazují žádné písmeno na prázdné slovo.

Důležité zobecnění těchto korespondencí se týká takzvaných pozitivních variet. Definici *pozitivní variety regulárních jazyků* získáme z definice 3.36 odstraněním požadavku uzavřenosti na komplementy. Dá se ukázat, že tyto variety odpovídají pseudovarietám konečných uspořádaných monoidů, přičemž příslušná korespondence se navazuje analogicky předchozí větě pomocí syntaktických uspořádaných monoidů.

3.4 Příklady

Pseudovarieta všech grup **G** je zadaná pseudoidentitou $x^\omega = 1$ a charakterizuje podle cvičení 1.19 právě varietu jazyků přijímaných duálně deterministickými automaty.

Cvičení 3.40. Ukažte, že pro varietu jazyků $\mathscr{L}$, která odpovídá pseudovarietě všech konečných komutativních grup, sestává $\mathscr{L}(A)$ právě z booleovských kombinací jazyků tvaru $\{w \in A^* \mid |w|_a \equiv k \pmod{n}\}$ pro $a \in A$, $k \in \mathbb{N}_0$ a $n \in \mathbb{N}$.

Cvičení 3.41. Ukažte, že pseudovarietě **N** všech nilpotentních pologrup, zadané pseudoidentitou $x^\omega = 0$, odpovídá v Eilenbergově korespondenci varieta $\mathscr{L}$ taková, že $\mathscr{L}(A)$ sestává právě z konečných jazyků a jazyků, jejichž komplement v A^+ je konečný.

Cvičení 3.42. Ukažte, že pseudovarietě **Sl** všech polosvazů, která je zadaná identitami $x^2 = x$ a $xy = yx$, odpovídá v Eilenbergově korespondenci varieta $\mathscr{L}$ taková, že $\mathscr{L}(A)$ sestává právě z booleovských kombinací jazyků A^*aA^* pro $a \in A$ (tato varieta tedy obsahuje právě jazyky, kde příslušnost slov je možné určit podle toho, jaká obsahují písmena).

Pseudovarieta všech aperiodických monoidů **A** (připomeňme, že podle tvrzení 2.70 se jedná právě o $\mathscr{H}$-triviální monoidy) je zadaná pseudoidentitou $x^{\omega+1} = x^{\omega}$ a odpovídá jí podle věty 2.71 varieta všech star-free jazyků.

Pro daný star-free jazyk je možné se ptát, kolikrát se v rozšířeném racionálním výrazu bez iterace, který tento jazyk zadává, musí vystřídat aplikace booleovských operací a zřetězení. Takto je definována *Straubingova–Thérienova zřetězovací hierarchie* star-free jazyků. Pro danou třídu regulárních jazyků $\mathscr{L}$ definujeme její *polynomiální uzávěr* $\mathrm{Pol}(\mathscr{L})$ tak, že pro každou abecedu A sestává $\mathrm{Pol}(\mathscr{L})(A)$ z konečných sjednocení jazyků tvaru $L_0 a_1 L_1 \cdots a_k L_k$, kde $k \in \mathbb{N}_0$, $a_i \in A$ a $L_i \in \mathscr{L}(A)$. Analogicky definujeme *booleovský uzávěr* $\mathrm{B}(\mathscr{L})$ třídy $\mathscr{L}$: množinu $\mathrm{B}(\mathscr{L})(A)$ tvoří právě jazyky získané z jazyků patřících do $\mathscr{L}(A)$ pomocí booleovských operací.

Je možné dokázat, že polynomiální uzávěr libovolné variety je pozitivní varieta.

Cvičení 3.43. Dokažte, že booleovský uzávěr pozitivní variety je varieta.

Cvičení 3.44. Dokažte, že pokud je $\mathscr{L}$ varieta a navíc $\mathscr{L}(A)$ obsahuje všechny jazyky $\{a\}$ pro $a \in A$, tak $\mathrm{Pol}(\mathscr{L})(A)$ sestává z konečných sjednocení jazyků tvaru $L_0 L_1 \ldots L_k$, kde $k \in \mathbb{N}_0$ a $L_i \in \mathscr{L}(A)$.

Operace na pseudovarietách monoidů odpovídající polynomiálnímu uzávěru je založena na důležitém pojmu Maľcevova součinu. Maľcevův součin pseudovariet se nejsnadněji definuje pomocí relačních homomorfismů monoidů.

Definice 3.45. Nechť S a T jsou monoidy. Relace $\sigma \subseteq S \times T$ se nazývá *relační homomorfismus*, jestliže splňuje následující podmínky:

1. $\forall x \in S\ \exists z \in T\colon (x,z) \in \sigma$.
2. $\forall x, y \in S\colon \sigma(x) \cdot \sigma(y) \subseteq \sigma(xy)$.
3. $(1,1) \in \sigma$.

Ekvivalentně můžeme říct, že σ je podmonoidem monoidu $S \times T$, jehož projekce na první složku je surjektivní. Ihned se vidí, že každý homomorfismus i každá inverze k surjektivnímu homomorfismu jsou relační homomorfismy a že složením relačních homomorfismů dostaneme opět relační homomorfismus.

Definice 3.46. Nechť **V** je pseudovarieta konečných monoidů a **W** je pseudovarieta konečných uspořádaných pologrup. *Maľcevovým součinem* $\mathbf{W} \circledM \mathbf{V}$ rozumíme třídu všech konečných uspořádaných monoidů $(S, \leq)$, pro které existuje relační homomorfismus $\sigma \subseteq S \times T$ takový, že $T \in \mathbf{V}$ a pro všechny idempotentní prvky $e \in E(T)$ náleží podpologrupa $\sigma^{-1}(e) \subseteq S$ do **W**.

Cvičení 3.47. Dokažte, že $\mathbf{W} \circledM \mathbf{V}$ je pseudovarieta uspořádaných monoidů.

Analogicky je možné definovat Maľcevův součin i v případě pseudovariet neuspořádaných pologrup a monoidů. Následující příklad ukazuje, jak lze pomocí této neuspořádané verze Maľcevova součinu rozložit pseudovarietu **J**.

Příklad 3.48. Ukážeme, že $\mathbf{J} = \mathbf{N} \circledcirc \mathbf{Sl}$.

Nechť nejprve S je libovolný konečný $\mathscr{J}$-triviální monoid. Definujeme kongruenci $\sim$ na S předpisem

$$x \sim y \quad \Longleftrightarrow \quad x^\omega = y^\omega.$$

Abychom ověřili, že $\sim$ je skutečně kongruencí monoidu S, stačí podle cvičení 1.9 ukázat, že se jedná o pravou i levou kongruenci, přičemž tyto dvě vlastnosti se ukáží symetricky. Předpokládejme tedy, že platí $x \sim y$, a dokažme $xz \sim yz$ pro libovolné $z \in S$. Opakovaným použitím první pseudoidentity z (3.1) dostaneme

$$(xz)^\omega = (xz)^\omega xz = (xz)^\omega x^\omega z.$$

Indukcí tedy obdržíme $(xz)^\omega = (xz)^\omega (x^\omega z)^\omega$. Opakovaným použitím druhé pseudoidentity z (3.1) můžeme výsledný prvek dále upravit jako $(xz)^\omega (x^\omega z)^\omega = (x^\omega z)^\omega$. Jelikož podobně platí $(yz)^\omega = (y^\omega z)^\omega$, dostáváme z předpokladu $x^\omega = y^\omega$ požadovanou rovnost $(xz)^\omega = (yz)^\omega$. Pologrupa $S/{\sim}$ je polosvaz, neboť její idempotence vyplývá přímo z definice kongruence $\sim$ a její komutativita je řečena druhou pseudoidentitou v (3.2). Přitom každá třída kongruence $\sim$ je nilpotentní, neboť pro ekvivalentní prvky x a y díky aperiodicitě platí $x^\omega y = y^\omega y = y^\omega = x^\omega$ a analogicky $yx^\omega = x^\omega$.

Nechť naopak existuje relační homomorfismus $\sigma \subseteq S \times T$, kde T je polosvaz a vzor každého prvku T je nilpotentní pologrupa. Ověříme platnost obou pseudoidentit v (3.2). Rovnost $x^{\omega+1} = x^\omega$ plyne z toho, že podpologrupa S generovaná libovolným prvkem je nilpotentní. Protože pro libovolné prvky $x, y \in S$ platí $\sigma(x)\sigma(y) \subseteq \sigma(xy)$ i $\sigma(y)\sigma(x) \subseteq \sigma(yx)$ a přitom $\sigma(x)\sigma(y) = \sigma(y)\sigma(x) \neq \emptyset$ díky komutativitě T, obsahují $\sigma(xy)$ a $\sigma(yx)$ nějaký společný prvek $e \in T$. Protože pologrupa T je idempotentní, platí v $\sigma^{-1}(e)$ rovnosti $(xy)^\omega = 0 = (yx)^\omega$.

Souvislost mezi polynomy a Maľcevovým součinem ukazuje následující lemma.

Lemma 3.49. *Nechť $L_0, \ldots, L_k$ jsou regulární jazyky nad abecedou A a $a_1, \ldots, a_k$ písmena z A. Označme L jazyk $L_0 a_1 L_1 \cdots a_k L_k$ a uvažujme syntaktické homomorfismy $\varphi_{L_i} \colon A^* \to M_{L_i}$ a $\varphi_L \colon A^* \to M_L$. Dále označme σ relační homomorfismus z M_L do $M_{L_0} \times \cdots \times M_{L_k}$ definovaný jako složení $(\varphi_{L_0}, \ldots, \varphi_{L_k}) \circ \varphi_L^{-1}$. Potom pro každý idempotentní prvek $e \in E(M_{L_0} \times \cdots \times M_{L_k})$ splňuje pologrupa $\sigma^{-1}(e)$ pseudoidentitu $x^\omega \leq x^\omega y x^\omega$.*

Důkaz. Pro dané prvky $x, y \in \sigma^{-1}(e)$ existují podle definice relace σ slova $u, v \in A^*$ splňující $\varphi_L(u) = x$, $\varphi_L(v) = y$ a $\varphi_{L_i}(u) = \varphi_{L_i}(v)$ pro $i = 0, \ldots, k$, přičemž $\varphi_{L_i}(u)$ je idempotentní prvek M_i. Nechť $n \in \mathbb{N}$ splňuje $n > k$ a současně platí $\varphi_L(u^n) = x^\omega$. Uvažujme libovolnou dvojici $(w, \bar{w}) \in \mathrm{C}_L(u^n)$. Jelikož platí $wu^n\bar{w} \in L = L_0 a_1 L_1 \cdots a_k L_k$, podle Dirichletova zásuvkového principu musí být alespoň jedna z n kopií slova u obsažena celá

uvnitř některého z jazyků L_i. Jelikož $\varphi_{L_i}(u) = \varphi_{L_i}(v)$ je idempotentní prvek M_i, je možné k této kopii u přidat libovolný počet kopií slov u a v, aniž by se cokoli změnilo na příslušnosti celého slova do jazyka L. Proto platí $wu^n vu^n \bar{w} \in L$, a tedy $(w, \bar{w}) \in \mathrm{C}_L(u^n vu^n)$, což jsme chtěli dokázat. □

Věta 3.50 (Jean-Éric Pin a Pascal Weil, 1997)**.** *Nechť* **V** *je libovolná pseudovarieta konečných monoidů a* $\mathscr{L}$ *jí odpovídající varieta regulárních jazyků. Potom pseudovarieta uspořádaných monoidů odpovídající pozitivní varietě* $\mathrm{Pol}(\mathscr{L})$ *je* $\mathrm{Mod}(x^\omega \leq x^\omega y x^\omega) \circledM \mathbf{V}$.

Straubingova–Thérienova hierarchie tvoří řetězec tříd jazyků

$$\mathscr{S}\mathscr{T}_0 \subseteq \mathscr{S}\mathscr{T}_{1/2} \subseteq \mathscr{S}\mathscr{T}_1 \subseteq \mathscr{S}\mathscr{T}_{3/2} \subseteq \cdots,$$

jejichž sjednocením je varieta všech star-free jazyků. Přitom každá celá úroveň je varieta jazyků a každá poloviční úroveň pozitivní varieta. Jednotlivé třídy jsou definovány následovně:

$$\begin{aligned} \mathscr{S}\mathscr{T}_0(A) &= \{\emptyset, A^*\}, \\ \mathscr{S}\mathscr{T}_{n+1/2} &= \mathrm{Pol}(\mathscr{S}\mathscr{T}_n), \\ \mathscr{S}\mathscr{T}_{n+1} &= \mathrm{B}(\mathscr{S}\mathscr{T}_{n+1/2}), \end{aligned}$$

pro všechna $n \in \mathbb{N}_0$. Je známo, že neomezené střídání aplikování booleovských operací a zřetězování je pro definování všech star-free jazyků nutné, takže tato hierarchie je skutečně nekonečná.

Úroveň $1/2$ Straubingovy–Thérienovy hierarchie obsahuje pro abecedu A právě jazyky, která vzniknou jako konečná sjednocení jazyků $A^* a_1 A^* \cdots a_k A^*$, pro $k \in \mathbb{N}_0$ a $a_i \in A$. Jelikož relace „být (roztroušeným) podslovem“ je dobré předuspořádání (viz příklad 4.7), jedná se právě o jazyky, které s každým svým slovem w obsahují i všechna slova, jichž je w podslovem. Snadno se nahlédne, že těmto jazykům odpovídá pozitivní pseudovarieta monoidů $\mathrm{Mod}(1 \leq x)$.

Z definice patří do úrovně 1 Straubingovy–Thérienovy hierarchie právě booleovské kombinace jazyků $A^* a_1 A^* \cdots a_k A^*$. Jedná se tedy právě o jazyky, které lze popsat určením konečně mnoha povolených množin podslov do jisté délky k, přičemž slovo patří do jazyka právě tehdy, když množina všech jeho podslov délky nejvýše k je rovna jedné ze zadaných množin. Jinými slovy, příslušnost slova do jazyka je určena již tím, jaká obsahuje podslova do jisté délky, přičemž tato délka může být pro každý jazyk jiná. Proto se těmto jazykům říká *po částech testovatelné*. Známý výsledek Imreho Simona z roku 1975 říká, že této varietě jazyků odpovídá pseudovarieta monoidů **J**.

Cvičení 3.51. Dokažte, že racionální výraz $A^* abA^*$ zadává jazyk po částech testovatelný nad abecedou $A = \{a, b\}$, ale nikoli nad abecedou $A = \{a, b, c\}$.

Je možné ukázat, že jazyky úrovně $3/2$ nad abecedou A jsou právě konečná sjednocení jazyků tvaru $A_0^* a_1 A_1^* \cdots a_k A_k^*$, kde $k \in \mathbb{N}_0$, $a_i \in A$ a $A_i \subseteq A$, tedy polynomy popsané větou 2.94.

Cvičení 3.52. Pomocí věty 2.94 dokažte, že varietě $\mathscr{S}\mathscr{T}_{3/2}$ odpovídá pseudovarieta monoidů zadaná nekonečným systémem pseudoidentit tvaru $x^\omega \leq x^\omega y x^\omega$, kde x a y jsou libovolná slova splňující $\mathrm{alph}(y) \subseteq \mathrm{alph}(x)$.

Otázka algoritmické charakterizace druhé úrovně Straubingovy–Thérienovy byla dlouho jedním z nejvýznamnějších otevřených problémů v teorii regulárních jazyků. Pozitivní odpověď dali až Thomas Place a Marc Zeitoun v roce 2014 a v následujících letech se jim podařilo dokázat rozhodnutelnost i pro úrovně 5/2 a 7/2.

Analogicky Straubingově–Thérienově hierarchii je možné definovat i hierarchii pro star-free jazyky bez prázdných slov. Jen je třeba mírně upravit definici polynomiálního uzávěru: $\mathrm{Pol}(\mathscr{L})(A)$ sestává z konečných sjednocení jazyků bez prázdných slov tvaru $u_0 L_1 u_1 \cdots L_k u_k$, kde $k \in \mathbb{N}_0$, $u_i \in A^*$ a $L_i \in \mathscr{L}(A)$. Nultá úroveň této hierarchie, nazývané tradičně *dot-depth hierarchie*, je tvořena pro abecedu A jazyky $\emptyset$ a A^+. Úroveň $1/2$ této hierarchie je tvořena konečnými sjednoceními jazyků $u_0 A^* u_1 \cdots A^* u_k$, kde $u_i \in A^*$, a odpovídá pseudovarietě pologrup $\mathrm{Mod}(x^\omega \leq x^\omega y x^\omega)$. Charakterizace pro úrovně 1 a $3/2$ jsou rovněž známy, jsou ale mnohem komplikovanější.

Dot-depth hierarchie je svázána se Straubingovou–Thérienovou hierarchií pomocí následujícího tvrzení. Před jeho formulováním je užitečné zavést dvě důležité operace na pseudovarietách.

Pro libovolnou pseudovarietu konečných monoidů $\mathbf{V}$ definujeme $\mathbf{LV}$ jako pseudovarietu pologrup obsahující právě konečné pologrupy, jejichž všechny lokální podmonoidy patří do $\mathbf{V}$. Jelikož idempotentní prvky jsou právě prvky tvaru x^ω, můžeme z množiny pseudoidentit definující $\mathbf{V}$ vytvořit množinu pseudoidentit definující $\mathbf{LV}$ tak, že uvážíme novou proměnnou x, za každou proměnnou y v původních pseudoidentitách dosadíme $x^\omega y x^\omega$ a místo termu 1 použijeme x^ω. Například označíme-li $\mathbf{T} = \mathrm{Mod}(x = 1)$ triviální pseudovarietu monoidů, potom platí $\mathbf{LT} = \mathrm{Mod}(x^\omega y x^\omega = x^\omega)$.

Pro pseudovarietu monoidů $\mathbf{V}$ a pseudovarietu pologrup $\mathbf{W}$ se $\mathbf{V} * \mathbf{W}$ definuje jako pseudovarieta pologrup generovaná polopřímými součiny $S \rtimes_\varphi T$, kde $S \in \mathbf{V}$ a $T \in \mathbf{W}$, přičemž požadujeme $\varphi(t)(1) = 1$ pro všechny $t \in T$.

Věta 3.53 (Howard Straubing, 1985)**.** *Je-li* $\mathbf{V}$ *pseudovarieta monoidů odpovídající nějaké nenulové úrovni Straubingovy–Thérienovy hierarchie a* $\mathbf{W}$ *pseudovarieta pologrup odpovídající stejné úrovni dot-depth hierarchie, potom platí* $\mathbf{V} = \mathbf{W} \cap \mathbf{M}$ *a* $\mathbf{W} = \mathbf{V} * \mathbf{LT}$.

Další studovanou hierarchií variet jazyků je takzvaná *grupová hierarchie*, která je definovaná stejně jako Straubingova–Thérienova hierarchie, jen za nultou úroveň bereme místo triviální pseudovariety monoidů pseudovarietu $\mathbf{G}$. Je známo, že každou

úroveň této hierarchie je také možné vyjádřit pomocí příslušné úrovně Straubingovy–Thérienovy hierarchie jako $\mathbf{V} * \mathbf{G}$. O grupové hierarchii se podařilo dokázat hluboké výsledky charakterizující úrovně $1/2$ a 1.

Úroveň $1/2$ odpovídá pseudovarietě $\mathrm{Mod}(1 \leq x^\omega)$ a jazyky, které obsahuje, lze popsat rovněž topologicky. Uvažujme na A^* ultrametriku definovanou stejně jako d na straně 69, pouze místo konečných monoidů berme jen konečné grupy. Všimněte si, že takto skutečně obdržíme na A^* metriku, neboť podle příkladu 3.5 bude vzdálenost libovolných dvou slov nenulová. Na rozdíl od metriky d ovšem nebude prostor A^* s touto metrikou diskrétní, neboť například posloupnost slov $a^{n!}$ konverguje k prázdnému slovu. Potom úroveň $1/2$ grupové hierarchie sestává právě z regulárních jazyků otevřených vzhledem k této ultrametrice. Všimněte si, že pseudoidentita $1 \leq x^\omega$ vlastně vystihuje, že s libovolným slovem uv musí v jazyce ležet pro dostatečně velké n i slova $uw^{n!}v$, neboť tato posloupnost k uv konverguje.

Úroveň 1 grupové hierarchie odpovídá pseudovarietě $\mathrm{Mod}((x^\omega y^\omega)^\omega = (y^\omega x^\omega)^\omega)$, která sestává právě z monoidů, v nichž idempotentní prvky generují $\mathscr{J}$-triviální monoid. Ekvivalentně se tyto monoidy dají charakterizovat podmínkou, že každá $\mathscr{L}$-třída a každá $\mathscr{R}$-třída obsahuje nejvýše jeden idempotentní prvek.

Uzávěr libovolné variety regulárních jazyků $\mathscr{L}$ na polynomy a booleovské operace, tedy vlastně sjednocení celé hierarchie, která má $\mathscr{L}$ za nultou úroveň, je možné rovněž popsat pomocí Maľcevova součinu:

Věta 3.54 (Howard Straubing, 1979)**.** *Nechť* $\mathbf{V}$ *je libovolná pseudovarieta konečných monoidů a* $\mathscr{L}$ *jí odpovídající varieta regulárních jazyků. Potom pseudovarieta monoidů odpovídající uzávěru* $\mathscr{L}$ *na polynomy a booleovské operace je* $\mathbf{A} \circledM \mathbf{V}$.

Všimněte si, že volbou $\mathbf{V} = \mathbf{T}$ dostaneme z této věty charakterizaci star-free jazyků pomocí aperiodických monoidů.

Další regulární jazyky, které byly intenzívně studovány, jsou ty, u nichž lze příslušnost slova určit již ze znalosti, jaké toto slovo obsahuje faktory určité délky. Jazyky definované pomocí podmínek tohoto typu obvykle tvoří variety jazyků bez prázdných slov.

Cvičení 3.55. Dokažte, že varietě $\mathscr{L}$ jazyků bez prázdných slov, kde $\mathscr{L}(A)$ obsahuje právě booleovské kombinace jazyků wA^*, pro $w \in A^+$, odpovídá pseudovarieta pologrup $\mathrm{Mod}(x^\omega y = x^\omega)$. Všimněte si, že $\mathscr{L}$ obsahuje právě jazyky, které lze zadat nějakou podmínkou na prefixy délky nejvýše n, pro nějaké $n \in \mathbb{N}$.

Podobně snadno lze charakterizovat jazyky popsatelné pomocí prefixů a sufixů:

Cvičení 3.56. Dokažte, že varietě $\mathscr{L}$ jazyků bez prázdných slov, kde $\mathscr{L}(A)$ obsahuje právě booleovské kombinace jazyků wA^* a A^*w, pro $w \in A^+$, odpovídá pseudovarieta $\mathbf{LT}$.

Daleko obtížnější ovšem je charakterizovat jazyky popsatelné pomocí faktorů. Všimněte si, že aby tyto jazyky tvořily varietu, musíme připustit současně i testování pomocí

prefixů a sufixů. Jazyk L nad abecedou A se nazývá *lokálně testovatelný*, jestliže je booleovskou kombinací jazyků A^*wA^*, wA^* a A^*w, pro $w \in A^+$.

Věta 3.57 (Janusz A. Brzozowski a Imre Simon, 1973; Robert McNaughton, 1974)**.** *Varietě všech lokálně testovatelných jazyků odpovídá pseudovarieta pologrup* **LSl**.

Zajímavou varietu jazyků definuje rovněž pseudovarieta všech $\mathscr{R}$-triviálních monoidů $\mathbf{R} = \mathrm{Mod}((xy)^\omega x = (xy)^\omega)$, kterou je možné podobně jako její podpseudovarietu **J** psát ve tvaru $\mathbf{R} = \mathrm{Mod}(x^\omega y = x^\omega)$ Ⓜ **Sl**. Do této variety totiž patří právě jazyky, které umíme popsat pomocí podslov, s tím, že oproti po částech testovatelným jazykům navíc umožníme určovat pořadí, ve kterém se podslova ve směru od leva do prava poprvé vyskytují. Tyto jazyky je možné rovněž definovat jako konečná (disjunktní) sjednocení jazyků tvaru $A_0^* a_1 A_1^* \cdots a_k A_k^*$, kde $k \in \mathbb{N}_0$, $a_i \in A$, $A_i \subseteq A$ a $a_i \notin A_{i-1}$.

Poznámka 3.58. Mnoho přirozeně definovaných tříd regulárních jazyků lze charakterizovat i dalšími způsoby, z nichž nejdůležitější je charakterizace pomocí fragmentů monadické predikátové logiky druhého řádu na konečných lineárně uspořádaných množinách. Charakterizace těchto tříd jazyků pomocí monoidů potom vlastně dává algoritmický postup jak určit, zda k dané formuli této logiky existuje ekvivalentní formule, která patří do tohoto fragmentu.

Přesněji, uvažovaná logika používá binární relační symbol $\leq$ a unární relační symbol R_a pro každé písmeno $a \in A$. Za modely bereme konečné lineárně uspořádané množiny, jejichž prvky reprezentují pozice písmen ve slově; tedy slovo délky n je reprezentováno n-prvkovou lineárně uspořádanou množinou. Přitom symbol $\leq$ odpovídá uspořádání pozic odleva doprava a symbol R_a určuje, zda na dané pozici ve slově je písmeno a. Oproti logice prvního řádu máme navíc k dispozici kvantifikování přes množiny prvků, tedy přes množiny pozic ve slově. Například jazyk aa^* lze popsat formulí

$$(\forall x\colon R_a(x)) \wedge \exists X\colon \mathrm{first} \in X \wedge \mathrm{last} \notin X \wedge (\forall x\colon x \in X \iff \mathrm{next}(x) \notin X),$$

kde operační symboly first, last a next určují první, poslední a následující pozici, přičemž tyto pozice lze snadno popsat formulí prvního řádu se symbolem $\leq$.

Julius Richard Büchi v roce 1960 dokázal, že jazyk je regulární právě tehdy, když jej můžeme popsat pomocí formule této logiky. Robert McNaughton a Seymour Papert v roce 1971 ukázali, že star-free jazyky jsou právě jazyky, které lze definovat formulí nepoužívající kvantifikování přes množiny, tedy formulí prvního řádu. Přitom Straubingova–Thérienova hierarchie odpovídá hierarchii formulí prvního řádu určené střídáním kvantifikátorů. Přesněji, jazyky úrovně $n + 1/2$, pro $n \in \mathbb{N}_0$, jsou právě jazyky popsatelné Σ_{n+1}-formulí.

Literatura

[1] Berstel, Jean. Transductions and context-free languages. Teubner, Stuttgart, 1979. Opravená verze prvních čtyř kapitol je dostupná online.

[2] de Luca, Aldo; Varricchio, Stefano. Finiteness and regularity in semigroups and formal languages. Springer, Berlin, 1999.

[3] Filakovský, Marek. Rozpoznatelné a racionální podmnožiny pologrup. Bakalářská práce, Masarykova univerzita, 2009.

[4] Grillet, Pierre Antoine. Semigroups: an introduction to the structure theory. Marcel Dekker, New York, 1995.

[5] Pin, Jean-Éric. Varieties of formal languages. Plenum Publishing Corp., New York, 1986.

[6] Pin, Jean-Éric. Mathematical foundations of automata theory. Aktuální verze (18. 2. 2022) dostupná online.

[7] Rozenberg, Grzegorz; Salomaa, Arto, ed. Handbook of formal languages. Springer, Berlin, 1997.

[8] Sakarovitch, Jacques. Elements of Automata Theory. Cambridge University Press, Cambridge, 2009.

Rejstřík

akce
  pologrupy na množině, 7
  pologrupy na pologrupě, 60
automat
  deterministický, 8
    konečný, 8
  minimální, 12
  nedeterministický, 15
    konečný, 15

band, 43
  rektangulární, 43

$\mathscr{D}$-třída
  regulární, 38
dělitelnost
  pologrup, 4

ekvivalence
  Greenova $\mathscr{D}$, 28
  Greenova $\mathscr{H}$, 28
  Greenova $\mathscr{J}$, 27
  Greenova $\mathscr{L}$, 27
  Greenova $\mathscr{R}$, 27

faktor, 4
  hlavní, 42
funkce
  sekvenční, 63

grupa
  Schützenbergerho, 29

hierarchie
  dot-depth, 81
  grupová, 81
  Straubingova–Thérienova, 78
hodnost, 24
homomorfismus
  relační, 78
  syntaktický, 11
  vyhodnocovací, 52

ideál, 27
  levý, 27
  pravý, 27
index, 26
inverze, 36
  grupová, 37
iterace
  Kleeneho, 15
    pozitivní, 14

jazyk, 4
  lokálně testovatelný, 83
  po částech testovatelný, 80
  regulární, 19
  star-free, 45

kongruence, 4
  levá, 8
  pravá, 8
  syntaktická, 11
kontext, 11
kvocient
  jazyka, 17
  Reesův, 41

les

faktorizační, 52
lokální divisor, 56

monoid
flip-flop, 58
stop, 23
syntaktický, 12
transformační
úplný, 5
U_1, 62
U_2, 58
monom, 55

nasycování, 7

operace
implicitní, 68
racionální, 15
termová, 69

předuspořádání
dobré, 86
lexikografické, 86
syntaktické, 11
levé, 11
pravé, 11
převodník, 21
sekvenční, 63
přijímání
deterministickým automatem, 8
nedeterministickým automatem, 15
perioda, 26
podgrupa, 25
podmnožina
lineární, 23
nevyhnutelná, 90
pololineární, 23
racionální, 14
rozpoznatelná, 6
souvislá, 24
ultimativně periodická, 7
podmonoid
lokální, 55
podslovo, 4
pologrupa
0-jednoduchá, 42
aperiodická, 45
$\mathscr{H}$-triviální, 46
$\mathscr{J}$-triviální, 28
jednoduchá, 41
nilpotentní, 67
nulová, 41
přechodová, 9
periodická, 26
$\mathscr{R}$-triviální, 62
Reesova maticová, 43
syntaktická, 11
uspořádaná, 9
syntaktická, 12
polynom, 55
posloupnost
špatná, 86
prefix, 4
princip
věnčitého součinu, 63
prvek
grupový, 25
idempotentní, 3
nulový, 4
regulární, 36
pseudoidentita, 72
pseudovarieta, 65
A, 78
G, 66
J, 74
LSl, 83
LT, 81
M, 68
N, 67
R, 74
Sl, 77
T, 81

relace
    nezávislosti, 23
    racionální, 21
    rozpoznatelná, 20
    závislosti, 23
reprezentace
    Schützenbergerho, 44
rozpoznávání
    homomorfismem, 5
    pologrupou, 5
    uspořádanou pologrupou, 10
rozpoznatelnost, 6

součin
    kaskádový, 58
    Maľcevův, 78
    polopřímý, 60
    věnčitý, 61
splňovat
    pseudoidentitu, 72
stopa, 24
    souvislá, 24
sufix, 4

uspořádání
    monotónní, 9
uzávěr
    booleovský, 78
    polynomiální, 78

výška
    faktorizačního lesa, 52
výraz
    racionální, 14
varieta
    regulárních jazyků, 75
        pozitivní, 77
vlastnost
    konečné mocniny, 49

wqo, 86