Vysoké učení technické v Brně
Fakulta informačních technologií
Ústav informačních systémů

ZÁVĚREČNÁ PRÁCE

Ing. Martin Heckel

Dvourozměrné jazyky: dláždící (tiling) systémy

Brno, prosinec 2002

Abstrakt

Práce se zabývá zobecněním pojmů a technik teorie formálních jazyků z klasického jednorozměrného přístupu do dvou dimenzí. Běžně se dvourozměrné řetězce nazývají obrazy a jsou definovány jako pravoúhlé pole tvořené symboly nad jistou konečnou abecedou. Dvourozměrný jazyk (někdy nazývaný obrazový jazyk) je množina obrazů (dvourozměrných řetězců).

Tato práce navazuje na práci Ing. Vladimíra Bartíka: *Dvourozměrné jazyky a jejich formální prostředky* a úžeji se zaměřuje na tzv. dláždící (tiling) systémy, které jsou založené na projekcích lokálních obrazových jazyků (LOC). LOC je způsob definice obrazového jazyka pomocí množiny "dlaždic" (bloků 2 x 2). V práci jsou také diskutovány uzávěrové vlastnosti dláždících systémů (vzhledem k projekci, řádkové a sloupcové konkatenaci, řádkovému a sloupcovému uzávěru, sjednocení, průniku, komplementu a rotaci).

Klíčová slova

Lokální obrazový jazyk, dláždící (tiling) systém, uzávěrové vlastnosti.

Tato práce vznikla v rámci kurzu Moderní teoretická informatika, který v akademickém roce 2002/2003 vyučuje doc. RNDr. Alexander Meduna, CSc.

.....
Ing. Martin Heckel

Obsah

1. LOKÁLNÍ JAZYKY 4

1.1 LOKÁLNÍ JEDNOROZMĚRNÉ JAZYKY 4
1.2 LOKÁLNÍ DVOUROZMĚRNÉ JAZYKY 4
1.3 PROCEDURÁLNÍ NÁHLED NA DVOUROZMĚRNÉ LOKÁLNÍ JAZYKY 4
1.4 ZOBECNĚNÍ LOKÁLNÍCH JAZYKŮ 5
1.5 PŘÍKLADY 5

2. DLÁŽDÍCÍ (TILING) SYSTÉMY 6

2.1 PŘÍKLADY 6

3. UZÁVĚROVÉ VLASTNOSTI 8

3.1 DŮKAZ UZAVŘENOSTI *L(TS)* VZHLEDEM K PROJEKCI 8
3.2 DŮKAZ UZAVŘENOSTI *L(TS)* VZHLEDEM KE SLOUPC. A ŘÁDKOVÉ KONKATENACI 8
3.3 DŮKAZ UZAVŘENOSTI *L(TS)* VZHLEDEM K SLOUPC. A ŘÁDKOVÉMU UZÁVĚRU 10
3.4 DŮKAZ UZAVŘENOSTI *L(TS)* VZHLEDEM KE SJEDNOCENÍ A PRŮNIKU 10
3.5 DŮKAZ NEUZAVŘENOSTI *L(TS)* VZHLEDEM KE KOMPLEMENTU 11
3.6 DŮKAZ UZAVŘENOSTI *L(TS)* VZHLEDEM K ROTACI 12

4. LITERATURA 12

1. Lokální jazyky

1.1 Lokální jednorozměrné jazyky

Neformálně bychom mohli lokální jednodimenzionální jazyky (jazyky řetězců) definovat následovně: Lokální jazyk je určen pomocí množiny řetězců délky 2. Taková množina obsahuje všechny možné povolené faktory délky 2 pro řetězce jazyka (včetně ohraničení řetězců pomocí některého symbolu (např. #), který není obsažen v abecedě jazyka).

1.2 Lokální dvourozměrné jazyky

Definice 1.1:

Je dán obraz p velikosti *(m, n)*, nechť $h \le m$, $k \le n$, pak označíme $B_{h,k}(p)$ množinu všech bloků (podobrazů) p velikosti *(h, k)*. Dlaždicí nazveme čtvercový obraz o velikosti *(2, 2)*.

Definice 1.2:

Dvoudimenzionální jazyk $L \subseteq \Gamma^{**}$ je lokální, když existuje konečná množina Θ dlaždic nad abecedou $\Gamma \cup \{\#\}$ taková, že $L = \left\{ p \in \Gamma^{**} \middle| B_{2,2}(\overline{p}) \subseteq \Theta \right\}$. Γ je konečná abeceda. (Jazyk obsahuje prázdný obraz λ (o velikosti *(0, 0)*), jen když Θ obsahuje dlaždici $\begin{pmatrix} \# & \# \\ \# & \# \end{pmatrix}$.)

Θ tedy reprezentuje množinu povolených bloků obrazů patřících do lokálního jazyka L.

Jazyk L je lokální, můžeme-li jej přesně z množiny Θ vybrat. Množinu Θ nazýváme dlaždicovou reprezentací lokálního jazyka L a zapisujeme $L = L(\Theta)$. Třídu lokálních obrazových jazyků označíme *LOC*.

1.3 Procedurální náhled na dvourozměrné lokální jazyky

Na dvourozměrné lokální jazyky můžeme nahlížet také z pozice procedury, která rozeznává obraz: Okno o velikosti 2 × 2 se pohybuje po zkoumaném obrazu a při této

činnosti jsou pořizovány záznamy (bloky 2 × 2, které jsou oknem vidět) s každou změnou pozice okna, přičemž neuvažujeme jejich pořadí a četnost výskytů. Po skončení pohybu okna získáme množinu záznamů, která nám určí, zda je obraz přijat (množina záznamů je podmnožinou Θ), nebo nepřijat (množina záznamů není podmnožinou Θ).

1.4 Zobecnění lokálních jazyků

Podobně jako u jednorozměrného případu, můžeme i u dvourozměrných lokálních jazyků zobecnit velikost faktorů (dlaždic) na *(h, k)*. Tedy definici lokálních jazyků lze zobecnit na *(h, k)* lokální. Testujeme, zda podmnožina $B_{h,k}(p)$ zkoumaného obrazu p náleží do množiny Θ' povolených bloků o velikosti *(h, k)*.

Z pohledu této zobecněné definice nahlížíme na třídu jazyků *LOC* jako na třídu *(2, 2)* lokálních jazyků.

1.5 Příklady

Příkladem lokálního obrazového jazyka je např. jazyk obsahující všechny čtvercové obrazy, které mají na hlavní diagonále symbol 1, mimo hlavní diagonálu obsahují symbol 0. Takový lokální jazyk je určen abecedou Γ a množinou dlaždic Θ následovně:

$$\Gamma = \{0,1\},$$

$$\Theta = \left\{ \begin{array}{l} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \# & 1 \\ \# & 0 \end{pmatrix}, \begin{pmatrix} \# & 0 \\ \# & 0 \end{pmatrix}, \\ \begin{pmatrix} 0 & 0 \\ \# & \# \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ \# & \# \end{pmatrix}, \begin{pmatrix} 0 & \# \\ 1 & \# \end{pmatrix}, \begin{pmatrix} 0 & \# \\ 0 & \# \end{pmatrix}, \begin{pmatrix} \# & \# \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \# & \# \\ 1 & 0 \end{pmatrix}, \\ \begin{pmatrix} \# & \# \\ \# & 1 \end{pmatrix}, \begin{pmatrix} \# & \# \\ 0 & \# \end{pmatrix}, \begin{pmatrix} \# & 0 \\ \# & \# \end{pmatrix}, \begin{pmatrix} 1 & \# \\ \# & \# \end{pmatrix} \end{array} \right\}.$$

Příkladem jazyka, který není lokální může být jazyk, který obsahuje čtvercové obrazy nad abecedou s jedním symbolem. Není zde žádná „lokální strategie“, jak porovnat počet řádků a sloupců.

2. Dláždící (tiling) systémy

Na základě definice lokálních jazyků a projekce můžeme definovat dláždící systémy, které jsou dalším nástrojem pro specifikaci obrazových jazyků.

Definice 2.1:

Dláždící systém je čtveřice $T = (\Sigma, \Gamma, \Theta, \pi)$, kde:
Σ, Γ jsou konečné abecedy,
Θ je konečná množina dlaždic nad abecedou $\Gamma \cup \{\#\}$,
$\pi : \Gamma \to \Sigma$ je projekce.

(Nechť $p \in \Gamma^{**}$. Potom projekce při zobrazení π obrazu p je obraz $p' \in \Sigma^{**}$ takový, že $p'(i, j\} = \pi(p(i, j))$ pro všechna $1 \le i \le l_1(p)$, $1 \le j \le l_2(p)$.)
(Nechť $L \subseteq \Gamma^{**}$. Potom projekce při zobrazení π jazyka L je jazyk $L' = \{p' \mid p' = \pi(p)\ \forall p \in L\} \subseteq \Sigma^{**}$.)

Dláždící systém T definuje jazyk L na abecedou Σ následovně: $L = \pi(L')$, přičemž L' je základní lokální jazyk nad abecedou Γ, který koresponduje s množinou dlaždic Θ, $L' = L(\Theta)$.

Jazyk $L \subseteq \Gamma^{**}$ je jazyk přijímaný dláždícím systémem, jestliže existuje dláždící systém $T = (\Sigma, \Gamma, \Theta, \pi)$ takový, že $L = L(T)$. Množinu všech takových dvourozměrných jazyků označujeme *L(TS)* a nazýváme jí třídou jazyků přijímaných dláždícími systémy. (Jazyk *L* patří do *L(TS)*, je-li projekcí některého lokálního jazyka.)

2.1 Příklady

Příkladem jazyka, který patří do třídy *L(TS)* může být jazyk definovaný následovně:

$\Sigma = \{a\}$,
$L = \{x \mid l_1(x) = l_2(x)\} \subseteq \Sigma^{**}$ - jazyk čtverců nad nad abecedou Σ.

Můžeme jednoduše ukázat, že takový jazyk je přijímán dlaždícími systémy. Jako základní lokální jazyk L' použijeme jazyk čtverců nad abecedou $\Gamma = \{0,1\}$ s symboly 1 na hlavní diagonále a se symboly 0 na zbylých pozicích (viz kapitola 1.5). Poté aplikujeme projekci $\pi : \Gamma \to \Sigma$ takovou, že $\pi(1) = \pi(2) = a$. Je zřejmé, že $L = \pi(L')$.

Podobný postup bychom mohli použít k důkazu, že jazyk čtverců nad jakoukoliv abecedou je jazykem přijímaným dláždícími systémy. Například použijeme-li abecedu $\Sigma = \{a,b\}$, lokální jazyk nad abecedou $\Gamma = \{a_0, b_0, a_1, b_1\}$ bude takový, že na hlavní diagonále budou jen symboly a_1, b_1, zatímco zbývající pozice budou zaplněny symboly a_0, b_0. Potom můžeme použít projekci, která nám „rekonstruuje“ každé a_i na a každé b_i na b pro $i = 0, 1$.

Další příklad jazyka, který patří do třídy *L(TS)* je jazyk:

$\Sigma = \{a\}$

$L = \{p \mid p \in \Sigma^{**} \wedge l_2(p) = 3\}$ - jazyk obsahující obrazy se 3 sloupci

Lokální jazyk L' je v tomto případě definován množinou dlaždic Θ (nad abecedou $\Gamma \cup \{\#\}$, $\Gamma = \{1,2,3\}$) následujícího obrazu:

$$\begin{pmatrix} \# & \# & \# & \# & \# \\ \# & 1 & 2 & 3 & \# \\ \# & 1 & 2 & 3 & \# \\ \# & 1 & 2 & 3 & \# \\ \# & 1 & 2 & 3 & \# \\ \# & 1 & 2 & 3 & \# \\ \# & 1 & 2 & 3 & \# \\ \# & 1 & 2 & 3 & \# \\ \# & \# & \# & \# & \# \end{pmatrix}.$$

Projekce $\pi : \Gamma \to \Sigma$ je definována $\pi(1) = \pi(2) = \pi(3) = a$.

Podobně lze ukázat, že jazyk $L = \{x \mid l_2(x) = c\} \subseteq \Sigma^{**}$, kde c je jakékoliv kladné celé číslo, patří do třídy *L(TS)*. Pro základní lokální jazyk postačuje abeceda s c symboly a množina Θ je tvořena dlaždicemi z ohraničeného (pomocí symbolů #) obrazu s c sloupci, kdy každý takový sloupec je tvořen jedním symbolem z této abecedy.

3. Uzávěrové vlastnosti

Třída *L(TS)*, třída jazyků přijímaných dlaždícími systémy, je uzavřena vzhledem k:

1) projekci,
2) řádkové konkatenaci,
3) sloupcové konkatenaci,
4) řádkovému uzávěru,
5) sloupcovému uzávěru,
6) sjednocení,
7) průniku,
8) rotaci.

Třída *L(TS)* není uzavřena vzhledem ke komplementu.

3.1 Důkaz uzavřenosti *L(TS)* vzhledem k projekci

Nechť jsou Σ_1, Σ_2 konečné abecedy a $\varphi : \Sigma_1 \to \Sigma_2$ - projekce. Musíme dokázat: Jestliže jazyk $L_1 \subseteq \Sigma_1^{**}$ je přijímán dláždícími systémy, potom $L_2 = \varphi(L_1)$ je také jazyk přijímaný těmito systémy.

Nechť $T_1 = (\Sigma_1, \Gamma, \Theta, \pi_1)$ je dláždící systém pro L_1, tedy $L_1 = L(T_1)$.
Potom $L_1 = \pi_1(L')$, kde L' je základní lokální jazyk reprezentovaný Θ.
Jazyk L' je také ovšem základním lokálním jazykem pro L_2.
$L_2 = \pi_2(L')$, kde $\pi_2 = \varphi.\pi_1 : \Gamma \to \Sigma_2$.

Odtud plyne, že $T_2 = (\Sigma_2, \Gamma, \Theta, \pi_2)$ je dláždící systém pro jazyk L_2.

3.2 Důkaz uzavřenosti *L(TS)* vzhledem ke sloupc. a řádkové konkatenaci

Nechť L_1, L_2 jsou obrazové jazyky nad abecedou Σ a jazyk L je jazyk odpovídající sloupcové konkatenaci $L_1 \phi L_2$. Z definice sloupcové konkatenace plyne, že obraz $p \in L$ je

tvořen kompozicí dvou obrazů $p_1 \in L_1$, $p_2 \in L_2$ se stejným počtem řádků, přičemž nejpravější sloupec $p_1 \in L_1$ je připojen k nejlevějšímu sloupci $p_2 \in L_2$.

Nechť $(\Sigma, \Gamma_1, \Theta_1, \pi_1)$ a $(\Sigma, \Gamma_2, \Theta_2, \pi_2)$ jsou dláždící systémy pro jazyky L_1 a L_2. Předpokládejme bez újmy na obecnosti, že Γ_1 a Γ_2 jsou disjunktní.

Můžeme definovat dláždící systém $(\Sigma, \Gamma, \Theta, \pi)$ pro jazyk *L*:

$\Gamma = \Gamma_1 \cup \Gamma_2$,

$\Theta_1' = \left\{ \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \middle| \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \Theta_1 \wedge b_1, d_1 \neq \# \right\}$ - množina všech dlaždic z Θ_1, mimo těch obsahujících pravé ohraničení obrazů,

$\Theta_2' = \left\{ \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \middle| \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \in \Theta_2 \wedge a_2, c_2 \neq \# \right\}$ - množina všech dlaždic z Θ_2, mimo těch obsahujících levé ohraničení obrazů,

$$\Theta_{12} = \left\{ \begin{array}{l} \begin{pmatrix} a_1 & a_2 \\ \# & \# \end{pmatrix}, \begin{pmatrix} \# & \# \\ b_1 & b_2 \end{pmatrix}, \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \end{pmatrix} \Bigg| \\ \begin{pmatrix} a_1 & \# \\ \# & \# \end{pmatrix}, \begin{pmatrix} \# & \# \\ b_1 & \# \end{pmatrix}, \begin{pmatrix} c_1 & \# \\ d_1 & \# \end{pmatrix} \in \Theta_1, \begin{pmatrix} \# & a_2 \\ \# & \# \end{pmatrix}, \begin{pmatrix} \# & \# \\ \# & b_2 \end{pmatrix}, \begin{pmatrix} \# & c_2 \\ \# & d_2 \end{pmatrix} \in \Theta_2 \end{array} \right\}$$

- množina všech „spojujících“ dlaždic, které odpovídají dvěma krajním spojeným sloupcům,

$\Theta = \Theta_1' \cup \Theta_2' \cup \Theta_{12}$.

Projekce $\pi : \Gamma \to \Sigma$ je definována: $\forall a \in \Gamma, \quad \pi(a) = \begin{cases} \pi_1(a), \textit{ pokud } a \in \Gamma_1 \\ \pi_2(a), \textit{ pokud } a \in \Gamma_2 \end{cases}$.

Obdobně bychom dokázali uzavřenost třídy *L(TS)* vůči řádkové konkatenaci. $L_1 \theta L_2$. Z definice řádkové konkatenace plyne, že obraz $p \in L$ je tvořen kompozicí dvou obrazů $p_1 \in L_1$, $p_2 \in L_2$ se stejným počtem sloupců, přičemž nejspodnější řádek $p_1 \in L_1$ je připojen k nejvrchnějšímu řádku $p_2 \in L_2$. Postupujeme obdobně jako u předchozího důkazu, rozdíl bude pouze v definici množiny $\Theta = \Theta_1' \cup \Theta_2' \cup \Theta_{12}$, přičemž Θ_1' bude reprezentovat množinu všech dlaždic z Θ_1, mimo těch obsahujících dolní ohraničení

obrazů, $\Theta_2^{'}$ bude reprezentovat množinu všech dlaždic z Θ_2, mimo těch obsahujících horní ohraničení obrazů a Θ_{12} bude množina všech „spojujících“ dlaždic, které odpovídají dvěma krajním spojeným řádkům.

3.3 Důkaz uzavřenosti *L(TS)* vzhledem k sloupc. a řádkovému uzávěru

Nechť L je obrazový jazyk nad abecedou Σ, který je přijímán dláždícími systémy. Z definice sloupcového uzávěru $L^{*\phi}$ vyplývá, že jazyk L je dán posloupností sloupcových konkatenací mezi obrazy v L. Potom můžeme najít takový dláždící systém $(\Sigma,\Gamma,\Theta,\pi)$ pro $L^{*\phi}$, přičemž využijeme stejného postupu jako u předchozího důkazu (viz kapitola 3.2).

Uvažujme dva různé dláždící systémy pro jazyk L: $(\Sigma,\Gamma_1,\Theta_1,\pi_1)$ a $(\Sigma,\Gamma_2,\Theta_2,\pi_2)$ tak, že abecedy Γ_1 a Γ_2 jsou disjunktní. Definujeme $\Gamma = \Gamma_1 \cup \Gamma_2$ a sestrojíme množinu Θ_{12} jako v předchozím důkazu. Potom množina dlaždic Θ je definována následovně: $\Theta = \Theta_1 \cup \Theta_{12}$. Projekce $\pi : \Gamma \to \Sigma$ je opět definována stejně jako v předchozím důkazu.

Stejná myšlenka se použije i při konstrukci dláždícího systému pro obrazový jazyk $L^{*\theta}$ odpovídající řádkovému uzávěru L.

3.4 Důkaz uzavřenosti *L(TS)* vzhledem ke sjednocení a průniku

Nechť L_1 a L_2 jsou dva obrazové jazyky nad abecedou Σ a nechť $(\Sigma,\Gamma_1,\Theta_1,\pi_1)$ a $(\Sigma,\Gamma_2,\Theta_2,\pi_2)$ jsou dva dláždící systémy, které rozpoznávají jazyk L_1 a L_2. Předpokládáme, že abecedy Γ_1 a Γ_2

ERROR: undefined
OFFENDING COMMAND: eexec

STACK:

-filestream-
-mark-