# Csempe átíró nyelvtanok

## Tile rewriting grammars

Németh L. Zoltán

Számítástudomány Alapjai Tanszék
SZTE, Informatikai Tanszékcsoport

1. előadás - 2006. április 10.

# Képek (pictures) I.

## Alapdefiníciók

- **ábécé**: $\Sigma$: nemüres, véges halmaz
- **kép**: $p$: egy mátrix melynek elemei $\Sigma$-beliek
- **méret**: $|p| = (|p|_{row}, |p|_{col})$
- **pixel**: $p(i,j)$: az $i$. sor, $j$. eleme $p$ mátrixában
  $1 \leq i \leq |p|_{row}, 1 \leq j \leq |p|_{col}$
- **üres kép**: $\lambda$
- **nemüres képek halmaza**: $\Sigma^{+,+}$
- **képek halmaza**: $\Sigma^{*,*} = \Sigma^{+,+} \cup \{\lambda\}$
- **képnyelv** (picture language): $L \subseteq \Sigma^{*,*}$

# Képek II.

## Homogenitás

- Ha $\Sigma' \subseteq \Sigma$ és egy $p \in \Sigma^{*,*}$ minden pixelje eleme $\Sigma'$-nek, akkor azt mondjuk, hogy $p$ $\Sigma'$**-homogén.**
- speciálisan: $a$**-kép**, ha $\{a\}$-homogén, $(a \in \Sigma)$.
- A $p \in \Sigma^{+,+}$ kép **homogén**, ha $a$-kép, valamely $a \in \Sigma$-ra.

## Rögzített méretű képek

- $\Sigma^{h,k}$: a $h \times k$ méretű képek halmaza.
- $a^{h,k}$: a $\Sigma^{h,k}$-beli $a$-kép.
- $a^{+,+}$: a $\Sigma^{+,+}$-beli $a$-képek halmaza.

# Jelölések

## Szokásos konvenciók

- képek: $p, q, r, \ldots$
- pozíciók (koordináta párok): $(i, j)$, $(i_1, j_1)$, $\ldots$
- ábécék: $\Sigma$, $\Delta$, $\Gamma$, $\ldots$
- betűk:

  $a$, $b$, $c$, $\ldots$ (terminálisok)

  $A$, $B$, $C$, $\ldots$, $X$, $Y$, $Z$, $S$, $\ldots$ (nemterminálisok)
- kezdőszimbólum: $S$

# Műveletek képeken és képnyelveken

## Műveletek

- **vízszintes konkatenáció**: $p \varobar q$ csak, ha $|p|_{row} = |q|_{row}$
- **függőleges konkatenáció**: $p \ominus q$ csak, ha $|p|_{col} = |q|_{col}$
- $\varobar$ és $\ominus$ kiterjeszthető képnyelvekre.
- $\varobar$ és $\ominus$ **lezárási operátor**ai:

$$L^{*\varobar} := \bigcup_{i \geq 0} L^{i\varobar}, \text{ ahol}$$

$$L^{0\varobar} := \{\lambda\}, \quad L^{i\varobar} = L \varobar L^{(i-1)\varobar}, \quad i > 0$$

- $L^{*\ominus}$ hasonlóan definiálható
- $L^{*,*}$ **Simplot-operátor**
(partícionálás segítségével definiálható).

# Részképek, részkép koordinátái

## Részképek

- $q \trianglelefteq_{(i,j)} p$: $q$ a $p$ **részképe** az $(i,j)$ (balfelső) pozícióban, ha

  $q(x,y) = p(i{-}1{+}x, j{-}1{+}y)\ \forall 1 \leq x \leq |q|_{row}, \forall 1 \leq y \leq |q|_{col}$
- $q \trianglelefteq p$: ha létezik $(i,j)$, hogy $q \trianglelefteq_{(i,j)} p$.

## Koordináták

- $q$ **koordinátái** $p$-ben:

  $coor_{(i,j)}(q,p) := \{ (x,y) | i \leq x < i{+}|q|_{row}, j \leq y < j{+}|q|_{col} \}$
- $coor_{(i,j)}(q,p) = \emptyset$, ha $q \trianglelefteq p$ nem teljesül
- $coor(p) = coor_{(1,1)}(p,p)$

# Partícionálás I.

## Partícionálás

A $\Pi(p) = \{\, (p_1, i_1, j_1), \ldots, (p_n, i_n, j_n) \,\}$ halmaz a $p$ kép egy partícionálása, ha

1. $p_t \trianglelefteq p$, $\forall t : 1 \leq t \leq n$-re, és
2. $\{coor_{(i_1, j_1)}(p_1), \ldots, coor_{(i_n, j_n)}(p_n)\}$ a $coor(p)$ partícionálása.

## Simplot-operátor: $L^{*,*}$

- $p \in L^{+,+} \Leftrightarrow p$ $L$-beli részképekre particionálható.
- $L^{*,*} := L^{+,+} \cup \{\lambda\}$.
- $L^{*,*}$ bővebb a $\varobar$ és $\ominus$-ra egyszerre vonatkozó lezártnál.

# Partícionálás II.

## Homogén partícionálás

A $\Pi(p)$ partícionálás **homogén**, ha benne minden részkép homogén (azaz csak azonos betűkből áll.)

## $L$ partícionálás halmaza

Legyen $L$ képnyelv, egy $\Pi = \{(p, \Pi(p)) \mid p \in L\}$ halmaz $L$ **(homogén) partícionálás halmaza**, ha minden $\Pi(p)$ (homogén) partícionálása $p$-nek.

# Helyettesítés fogalmak

## Helyettesítés leképezés megadása

Legyen $\Sigma$ és $\Delta$ két ábécé. Egy $\sigma : \Delta \to 2^{\Sigma^{+,+}}$ leképezést (képnyelv) **helyettesítésnek** hívunk ($\Delta$-ból $\Sigma$-ba).

## Helyettesítés (egy konkrét részkép helyettesítése)

Legyenek $p$, $q$, $q'$ képek, melykre $|q| = |q'|$ és $q \trianglelefteq_{(i,j)} p$ teljesül. Ekkor $p[q'/q]_{(i,j)}$ jelölje azt a képet, melyet úgy kapunk, hogy $p$-ben a $q$ részkép $(i,j)$ pozíción való előfordulását a $q'$ képre cseréljük, azaz

$$p[q'/q]_{(i,j)}(i-1+x, j-1+y) = q'(x,y)$$

$$\forall 1 \leq x \leq |q|_{row}, \forall 1 \leq y \leq |q|_{col}.$$

# Helyettesítés adott partícionálás szerint

## Helyettesítés képben homogén partícionálás szerint

Legyen
$\sigma : \Delta \to 2^{\Sigma^{+,+}}$ helyettesítés
$p \in \Delta^{+,+}$ egy kép
$\Pi(p) = \{(p_1, i_1, j_1), \ldots, (p_n, i_n, j_n)\}$ egy homogén partícionálás, ahol minden $p_t$ egy $d_t$-kép, valamely $d_t \in \Delta$-ra, $\forall 1 \leq t \leq n$. Ekkor a $p$-ben $\Pi(p)$ **által indukált helyettesítés**:

$$\sigma_\Pi(p) := \{p[r_1/p_1]_{(i_1,j_1)} \ldots [r_n/p_n]_{(i_n,j_n)} \mid r_t \in \sigma(d_t), 1 \leq t \leq n\}$$

## Helyettesítés nyelvben homogén partícionálás halmaz szerint

Legyen $L \subseteq \Delta^{+,+}$, $\Pi$ $L$ egy homogén partícionálás halmaza, ekkor az $L$-ben $\Pi$ **által indukált** $\sigma_\Pi$ **helyettesítés**e:

$$\sigma_\Pi(L) := \{\sigma_{\Pi(p)}(p) \mid p \in L\}.$$

# Részképek sarkainak megjelölése

## Blokk-képek

- Ha $p \in \Delta^{+,+}$ egy kép, legyen $\square(p)$ az a kép, amelyet $p$ sarkainak a megjelölésével kapunk.
- Pl.

$$p = \begin{pmatrix} b & b & b & b \\ b & b & b & b \end{pmatrix} \quad \square(p) = \begin{pmatrix} {}^1b & b & b & b^2 \\ {}_4b & b & b & b_3 \end{pmatrix}$$

$$q = \begin{pmatrix} a & a & a \end{pmatrix} \quad \square(q) = \begin{pmatrix} {}^1_4a & a & a^2_3 \end{pmatrix}$$

- $\square(p)$ megadható $\Delta \times \mathcal{M}$ feletti képként, ahol

$$\mathcal{M} = \{\,\cdot,\ {}^1,\ {}^2,\ {}_3,\ {}_4,\ {}^{12},\ {}_{34},\ {}^1_4,\ {}^2_3,\ {}^{12}_{34}\,\}$$

- $(\Delta \times \mathcal{M})^{+,+}$ elemeit **blokk-képeknek** hívjuk.

# Blokknyelvek

## Blokknyelvek

- Ha $L \subseteq \Delta^{+,+}$, akkor $\Box(L) := \{\Box(p) \mid p \in L\}$
- Ha $p \in \Box(d^{+,+})$ akkor $p$ **blokk-$d$-kép** másszóval **blokkhomogén** ($d \in \Delta$).
- **univerzális blokknyelv**:

$$\Delta^{\Box+,+} := \Big(\bigcup_{a \in \Delta} \Box(a^{+,+})\Big)^{+,+}$$

- $\Delta^{\Box+,+}$ felimerhető (csempéző rendszerekkel (TS))
- **blokknyelv**: $L \subseteq \Delta^{\Box+,+}$
  Azaz a blokknyelvek esetén a képek egy-egy homogén partícionálása indexekkel jelölve van.

# Blokkhelyettesítés

## Blokkhelyettesítés egy blokk-képbe

Ha $\sigma : \Delta \to 2^{\Sigma^{+,+}}$ helyettesítés, $p \in \Delta^{\square +,+}$ blokk-kép, akkor legyen $\Pi'$ $p$ (egyértelmű) partícionálása, blokkhomogén részképekre. Ekkor $p$ **blokkhelyettesítése**:

$$\sigma_B(p) := \sigma_{\Pi'}(p).$$

## Blokkhelyettesítés nyelvbe

Ha $L \subseteq \Delta^{\square +,+}$ blokknylev, ekkor $L$ **blokkhelyettesítése**:

$$\sigma_B(L) := \{q \mid \exists p \in L, q \in \sigma_B(p)\}$$

# Konvexitás

## Konvexitás egy betűre nézve

Egy $p \in \Delta^{+,+}$ kép $y$**-konvex**, valamely $y \in \Delta$-ra, ha bármely $x \in \Delta, x \neq y$-ra $p$-ben nem fordul elő az alábbi $2 \times 2$-es csempék egyike sem:

$$\begin{matrix} y & y \\ x & y \end{matrix} \qquad \begin{matrix} x & y \\ y & y \end{matrix} \qquad \begin{matrix} y & x \\ y & y \end{matrix} \qquad \begin{matrix} y & y \\ y & x \end{matrix}$$

## Konvexitás halmazra nézve

Ha $\Delta' \subseteq \Delta$, akkor egy $p \in \Delta^{+,+}$ kép $\Delta'$**-konvex**, ha $y$-konvex, minden $y \in \Delta'$-re.

# Maximális $a$-részképek

## Maximális $a$-részkép

Legyen $a \in \Delta$, $p \in \Delta^{+,+}$. Egy $q$ $a$-kép **maximális $a$-részképe** $p$-nek, ha részképe, azaz $q \trianglelefteq_{i,j} p$ és miden $q' \trianglelefteq_{i',j'} p$ $a$-részképére $p$-nek vagy

$$coor_{(i',j')}(q',p) \cap coor_{(i,j)}(q,p) = \emptyset \text{ vagy}$$

$$coor_{(i',j')}(q',p) \subseteq coor_{(i,j)}(q,p)$$

## Állítás

Legyen $p \in \Delta^{+,+}$, $\Delta' \subseteq \Delta$ és $a \in \Delta'$. Ha $p$ $\Delta'$-konvex kép, akkor $p$ maximális, $a$-homogén részképei diszjunktak. Ezért $p$ maximális homogén képekre történő felbontásakor a $\Delta'$ feletti részképek egyértelműen meghatározottak.

# Képek szélének megjelölése # szimbólumokkal

## Bekeretezett képek

- Legyen $\# \notin \Sigma$ egy ún. **keret szimbólum.**
- $p \in \Sigma^{+,+}$ **bekeretezett változata**: $\bar{p}$ legyen az a kép, amelyet úgy kapunk, hogy $p$-t mindenütt egy-egy $\#$ szimbólummal vesszük körbe:

$$\bar{p} = \begin{array}{ccccc} \# & \# & \cdots & \# & \# \\ \# & p(1,1) & \cdots & p(1,|p|_{col}) & \# \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \# & p(|p|_{row},1) & \cdots & p(|p|_{row},|p|_{col}) & \# \\ \# & \# & \cdots & \# & \# \end{array}$$

# Csempék és lokális nyelv definíciója

## Lokális nyelvek

- A $\Sigma$ feletti $(h, k)$ méretű képek halamazát jelölje $\Sigma^{h,k}$.
- Egy $p \in \Sigma^{+,+}$ kép $(h, k)$ méretű részképeinek a halmaza:

$$B_{h,k}(p) := \{q \in \Sigma^{h,k} : q \trianglelefteq p\}.$$

- Kiterjesztése nyelvekre: $B_{h,k}(L) := \{B_{h,k}(p) \mid p \in L\}$
- Ha $\theta \subseteq (\Sigma \cup \{\#\})^{(2,2)}$, akkor $\theta$ elemeit **csempék**nek (tiles) is hívjuk.
- $LOC(\theta) := \{\, p \in \Sigma^{+,+} \mid B_{2,2}(\bar{p}) \subseteq \theta\,\}$
- Az $L \subseteq \Sigma^{+,+}$ nyelv **lokális**, ha $L = LOC(\theta)$ valamely $\theta \subseteq (\Sigma \cup \{\#\})^{2,2}$ csempe halmazra.

# Csempéző rendszerek

### Csempéző rendszerek (Tiling Systems)

Egy **csempéző rendszer** (Tiling System, TS) a következő komponensekből áll $T = (\Sigma, \Gamma, \theta, \pi)$, ahol

- $\Sigma$ és $\Gamma$ két ábécé,
- $\pi : \Gamma \to \Sigma$ egy leképezés,
- $\theta$ $2 \times 2$-es csempék halmaza $\Gamma \cup \{\#\}$ felett (nyilván véges).

### A $T$ által felismert v. definiált nyelv

$$L(T) := \pi(LOC(\theta)).$$

### TS nyelvek

A csempéző rendszer által felismerhető nyelveket, **TS-felismerhető**, vagy röviden **TS-nyelv**eknek hívjuk.

# Korlátozott lokalitás

## Korlátozott lokális nyelv

Egy $L \subseteq \Sigma^{+,+}$ **korlátozott lokális nyelv** (restricted local language), ha létezik olyan $\Gamma$ (# nélkül!) feletti $\theta$ csempe halmaz, melyre $p \in L \Leftrightarrow$ ha a következők teljesülnek:

1. ha $|p|_{row} \geq 2, |p|_{col} \geq 2$, akkor $B_{2,2}(p) = \theta$
2. ha $|p|_{row} = 1, |p|_{col} \geq 2$, akkor $B_{1,2}(p) = \theta$
3. ha $|p|_{row} \geq 2, |p|_{col} = 1$, akkor $B_{2,1}(p) = \theta$

## Megjegyzés

A korlátozott lokális nyelvek a lokális nyelvek valódi részosztályát alkotják. De a TS nyelveket ekvivalens módon lehet korlátozott lokális nyelvek véges uniójának projekcióiként definiálni.

# Csempe átíró nyelvtanok (végre!)

### Véges nyelv

Jelölje $FIN(\Sigma)$ a $\Sigma$ feletti véges nyelvek halmazát.

### Konvex & korlátozott lokális

$LCVX_{\Delta'}(\Delta)$, pedig a $\Delta'$ konvex, korlátozott lokális nyelvek halmazát $\Delta$ felett, $\Delta' \subseteq \Delta$.

### Csempe átíró nyelvtan (Tile Rewriting Grammar, TRG)

Csempe átíró nyelvtan egy $G = (\Sigma, N, S, R)$ négyes, ahol

- $\Sigma$ a terminális ábécé
- $N$ a nemterminális ábácé
- $S$ a kezdőszimbólum
- $R \subseteq N \times (LCVX_N(N \cup \Sigma) \cup FIN(\Sigma))$ szabályok véges halmaza.

# Nemterminális részképek blokkolása

## $N$-blokkolás

Legyen $N$ egy véges halmaz. $N$-blokkolásnak nevezzük azt a

$$\square_N : (\Sigma \cup N)^{+,+} \to \big(\Sigma \cup (N \times \mathcal{M})\big)^{+,+}$$

leképezést, mely bármely $N$-konvex kép minden maximális $A$-homogén $q$ részképét $\square(q)$-val helyettesíti, minden $A \in N$-re.

# Levezetés

## Levezetés

Legyen $G = (\Sigma, N, S, R)$ egy TRG. A $G$ szerinti egy lépésben történő levezetés a következő reláció:

$$\Rightarrow_G \subseteq \left(\Sigma \cup (N \times \mathcal{M})\right)^{+,+} \times \left(\Sigma \cup (N \times \mathcal{M})\right)^{+,+},$$

melyre $p \Rightarrow_G p'$ akkor és csak akkor, ha

- $|p| = |p'|$
- $\exists A \to \Omega \in R$, melyre
- $\exists r \trianglelefteq_{(i,j)} p$, $r$ blokk $A$-homogén kép
- $\exists \omega \in \Omega$, $|\omega| = |r|$, hogy
- $p' = p[\Box_N(\omega)/r]_{(i,j)}$.

# Csempéző nyelvtanok kifejező ereje

## $G$ által generált nyelv

A $G = (\Sigma, N, S, R)$ TRG **által generált nyelv**

$$L(G) := \{\, p \in \Sigma^{+,+} \mid \Box(S^{|p|}) \Rightarrow_G^* p \,\}$$

## TRG nyelvek

Jelölje **TRG** a csempe átíró nyelvtannal generálható nyelvek osztályát,

## Tétel

$$TS \subsetneq TRG$$

# Normálforma tétel

### Terminális normálforma

Egy $G = (\Sigma, N, S, R)$ nyelvtan **terminális normálformájú**, ha szabályaiban terminálisok, csak $A \rightarrow \sigma$ alakban fordulnak elő, azaz csak $1 \times 1$-es véges képekben, a korlátozott lokális nyelvekben nem ($\sigma \in \Sigma$).

### Tétel

Minden $G$ TRG nyelvtanhoz létezik vele ekvivalens terminális normálformájú TRG.

# TRG zártsági tulajdonságai

## Tétel

A *TRG* nyelvosztály zárt a következő műveletekre

1. unió
2. sor / oszlop konkatenáció: $\oplus$ / $\ominus$
3. sor / oszlop konkatenáció lezártja: $*\oplus$ / $*\ominus$
4. projekció
5. elforgatás, tükrözés

## A bizonyításról

A szavakból álló környezetfüggetlen nyelvekre működő bizonyítások közvetlenül általánosíthatók.

# Nemrekurzív nyelvtanok

## Tétel

$$TS \subsetneq TRG$$

## Nemrekurzív TRG

Egy $G = (\Sigma, N, S, R)$ TRG-t **nemrekurzívnak** mondunk, ha

$$A \Rightarrow_G^* p$$

esetén $p$-ben nincs blokk-$A$-homogén részkép ($A \in N$).
Jelölje $nrTRG$ a nemrekurzív TRG-kel generálható képnyelvek osztályát.

## Tétel

$$TS = nrTRG$$

# Irodalomjegyzék


[1] A. Cherubini, S. Crespi Reghizzi, M. Pradella and P. San Pietro, Picture Languages: Tiling Systems versus Tile Rewriting Grammars, *Theor. Comput. Sci*, to appear.

[2] D. Giammaresi and A. Restivo, Two-Dimensional Languages, in: A. Saloma and G. Rozenberg (eds.), *Handbook of Formal Languages*, vol. 3, Springer-Verlag, Berlin, 1997, 215–267.

[3] O. Matz, Regular Expressions and Context-Free Grammars for Picture Languages, in: *Proc. of 14th Annu. Symp. on Theor. Aspects of Comp. Sci.*, LNCS **1200**, Springer-Verlag, Berlin, 1997, 283–294.

[4] S. Crespi Reghizzi, M. Pradella, Tile Rewriting Grammars, in: *Proc. of Seventh Internat. Conf. on Developments in Language Theory (DLT 2003)*, LNCS **2710**, Springer-Verlag, Berlin, 2003, 206–217.

[5] S. Crespi Reghizzi, M. Pradella, Tile Rewriting Grammars and Picture Languages, *Theor. Comp. Sci* , **340(2)**, 2005, 257–272.