

Reguláris szűkítésű determinisztikus top-down
fatranszformátorok
Tudományos tevékenység összefoglalása
PhD tézisek

Vágvölgyi Sándor
JATE Alkalmazott Informatikai Tanszék
Szeged Árpád tér 2

1995. szeptember 21.



1 Fatranszformáció osztályok kompozíciós hatványai által alkotott véges és végtelen hierarhiák

Rounds [10] és Thatcher [13] vezette be a top-down fatranszformátor fogalmát (az indukált fatranszformátor osztályt T -vel jelöljük) a szintakszis vezérelt fordítás [6] modeljeként. Rögtön bizonyítást nyert hogy a top-down fatranszformációk nem zártak a kompozícióra, azaz, $T \subset T^2$. Engelfriet [2] igazolta hogy a T kompozíciós hatványai valódi hierarhiát alkotnak, azaz $T^n \subset T^{n+1}$, $n \geq 1$. Érdekes módon, a determinisztikus top-down fatranszformációk osztálya (DT) szintén nem zárt a kompozícióra, azaz $DT \subset DT^2$, de [3] publikációban megmutattuk hogy $DT^2 = DT^3$. Tehát, a determinisztikus esetben, a DT^n ($n \geq 1$) hierarhia összeesik a második szintre.

Engelfriet [1] vezette be a reguláris szűkítésű top-down (determinisztikus és nemdeterminisztikus) fatranszformátorok fogalmát (az indukált fatranszformáció osztályokat DT^R és T^R jelöli). A reguláris szűkítés azt jelenti hogy az átírási szabályokhoz reguláris fanyelveket csatolunk, és egy szabály csak akkor alkalmazható az input fa egy szögpontjánál ha a közvetlen részfák benne vannak a szabályban megadott fanyelvekben.

Az \mathcal{A} reguláris szűkítésű fatranszformátorról azt mondjuk hogy determinisztikus, ha az alábbi feltétel teljesül. Ha két különböző átírási szabálynak megegyezik a bal oldala, akkor létezik olyan x_i változó hogy az x_i -nek megfelelő szűkítési halmazok diszjunktak. Nyilvánvaló hogy a fenti definíció alapján \mathcal{A} legfeljebb egy szabályt tud alkalmazni az input fa egy szögpontjánál. Az \mathcal{A} reguláris szűkítésű fatranszformátorról azt mondjuk hogy erősen determinisztikus, ha bármely két különböző átírási szabálynak különbözik a baloldala. Nyilvánvaló hogy a fenti definíció alapján \mathcal{A} determinisztikus is. A D_s prefix jelöli az "erősen determinisztikus" kifejezést. Ennek megfelelően például $D_s T^R$ jelöli a reguláris szűkítésű erősen determinisztikus top-down fatranszformátorok által indukált fatranszformációk osztályát. A [3] cikkben igazoltuk hogy

$$DFTA \circ DT = DT^2 = DT^3 = D_s T^{DR} = (D_s T^{DR})^2 \quad \leftarrow$$

továbbá a [4] publikációban megmutattuk hogy

$$D_s T^R \subset (D_s T^R)^2 = (D_s T^R)^3 = FTA \circ DT. \quad \leftarrow$$

Itt $(D)FTA$ a (determinisztikus) top-down faautomaták által indukált fatranszformációk osztálya, és $FTA \circ DT$ Rounds [10] transzformációs rendszere.

Engelfriet [1] megmutatta hogy DT^R zárt a kompozícióra, de T^R nem zárt, azaz $(DT^R)^2 = DT^R$ és $T^R \subset (T^R)^2$. Engelfriet [1], [2] eredményeinek a felhasználásával könnyen igazolható hogy $(T^R)^n$, $n \geq 0$ valódi hierarhia, azaz minden $n \geq 0$ indexre, $(T^R)^n \subset (T^R)^{n+1}$. Természetesen a kompozícióra való zártság miatt a megfelelő $(DT^R)^n$, $n \geq 1$ determinisztikus hierarhia összeesik az első szintre.

A determinisztikus reguláris szűkítésű, determinisztikus és nemdeterminisztikus top-down fatranszformátorokat [4] dolgozatban vezettük be és tanulmányoztuk. Az indukált

fatranszformáció osztályokat DT^{DR} és T^{DR} jelöli. A determinisztikus reguláris szűkítést determinisztikus top-down faautomatával lehet kiszámolni. Ismét könnyű igazolni hogy T^{DR} nem zárt a kompozícióra és a kompozíciós hatványok valódi hierarhiát alkotnak, azaz minden $n \geq 1$ indexre $(T^{DR})^n \subset (T^{DR})^{n+1}$. Valóban, a három hatvány hierarhia T^n ($n \geq 1$), $(T^R)^n$ ($n \geq 1$), $(T^{DR})^n$ ($n \geq 1$), mindegyik valódi, összerakható egy hierarhiába:

$$T^n \subseteq (T^{DR})^n \subseteq (T^R)^n \subseteq T^{n+1} \quad (n \geq 1)$$

amely végtelen, de még nyitott hogy mindegyik tartalmazás valódi-e.

A [4] értekezés fő irányvonala a determinisztikus reguláris szűkítésű determinisztikus fatranszformátorok és az általuk indukált transzformációk osztálya, DT^{DR} . Megmutattuk hogy DT^{DR} nem zárt a kompozícióra, de [4] munkában nyitva maradt hogy a kompozíciós hatványok valódi hierarhiát alkotnak-e.

Az [5] publikációban a reguláris szűkítésű nyelveket az alábbi módon iteráltuk anélkül hogy kiléptünk volna a reguláris erdők R osztályából. Legyen $DR_0 = DR$ és legyen minden $n \geq 1$ indexre DR_n a DR_{n-1} szűkítésű determinisztikus top-down faautomaták által felismert fanyelvek osztálya. A $DR_n \subseteq R$ tartalmazás teljesül minden $n \geq 0$ indexre. Továbbá $DR_{n-1} \subseteq DR_n$ minden $n \geq 1$ indexre. A fenti tartalmazás valóságát úgy bizonyítottuk be hogy megkonstruáltuk a következő fanyelveket.

Legyenek 1 és 0 nulla aritású szimbólumok, és legyen σ kettő aritású szimbólum. Azt mondjuk hogy egy fa páratlan (páros) ha páratlan (páros) sok 1 levelet tartalmaz. L_e (L_o) jelöli a páros (páratlan) fák halmazát. Megjegyezzük hogy $0 \in L_e$ és $1 \in L_o$. Minden $n \geq 0$ indexre, a C_n fanyelvet az alábbi módon definiáljuk:

- (a) $C_0 = \{1, 0\}$,
- (b) minden $n \geq 1$ indexre, C_n a legkisebb halmaz amely kielégíti az alábbi feltételeket.
 - (i) $1, 0 \in C_n$
 - (ii) $\sigma(t, r) \in C_n$ valahányszor $t \in C_{n-1}$ és $r \in C_n$.

A C_n halmaz elemeit n -szeresen egymásba ágyazott fésűknek nevezzük. Nyilvánvalóan minden $i < j$ indexekre, $C_i \subset C_j$. Legyen $C_n^e = C_n \cap L_e$ és $C_n^o = C_n \cap L_o$. Az [5] munkában igazoltuk hogy minden $n \geq 1$ indexre, $C_n^e \in DR_n - DR_{n-1}$. Továbbá $L_e \in R - \bigcup_{n=0}^{\infty} DR_n$. Tehát igazoltuk hogy ezek az osztályok valódi hierarhiát alkotnak a felismerhető fanyelvek osztályán belül, (amelyet R jelöl), azaz, minden $n \geq 0$ indexre, $DR_n \subset DR_{n+1}$. Továbbá igazoltuk hogy ez a hierarhia nem meríti ki a felismerhető erdők osztályát, azaz

$$\bigcup_{n=0}^{\infty} DR_n \subset R.$$

A [11] publikációban bebizonyítottuk hogy a $(DT^{DR})^n$ ($n \geq 1$) hierarhia valódi. Ugyan- ebben a munkában definiáltuk és tanulmányoztuk a DR_n szűkítésű determinisztikus top-down fatranszformátorokat (DT^{DR_n} jelöli az indukált fatranszformáció osztályt), és megmutattuk hogy minden $n \geq 0$ indexre

$$(DT^{DR})^{n+1} \subseteq DT^{DR_n}$$

amely szemléletesen azt jelenti hogy $n+1$ darab DR szűkítésű determinisztikus fatranszformátor kompozícióját kiszámolhatjuk egyetlen DR_n szűkítésű determinisztikus fatranszformátorral. A [12] publikációban megmutattuk az előbbi tartalmazás fordítottját, azaz, minden $n \geq 0$ indexre,

$$DT^{DR_n} \subseteq (DT^{DR})^{n+1} .$$

Ezzel bebizonyítottuk a

$$DT^{DR_n} = (DT^{DR})^{n+1} \quad (n \geq 0)$$

egyenlőséget. Továbbá azt kapjuk hogy DT^{DR_n} ($n \geq 0$) valódi hierarchiát alkot.

2 A DTA , DT , DT^{DR} , DT^R fatranszformáció osztályok kompozíciós monoidja

Tekintsük az $M = \{DTA, DT, DT^{DR}, DT^R\}$ fatranszformáció osztályok halmazát, és az alábbi két monoidot, amelyeket M révén definiálunk. Az M^* szabad monoidot (a konkatenáció műveletével), és az $[M] = \{Y_1 \circ \dots \circ Y_m \mid m \geq 0, Y_i \in M, 1 \leq i \leq m\}$, monoidot (a \circ kompozíció műveletével), amelyet M generál végesen.

Az M feletti szavak fatranszformáció osztályokat reprezentálnak a $\| \cdot \| : M^* \rightarrow [M]$ homomorfizmus révén. $\| \cdot \|$ -t az alábbi módon definiáljuk.

$$\|Y_1 \cdot Y_2 \cdot \dots \cdot Y_m\| = Y_1 \circ Y_2 \circ \dots \circ Y_m .$$

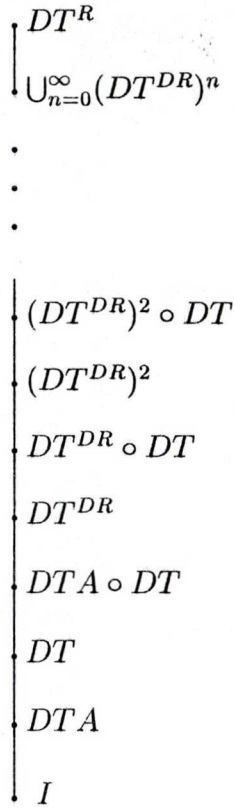
$I \in [M]$ jelöli az identitás fatranszformációkból álló osztályt, azaz, $I = \|\lambda\|$. Legyen θ a $\| \cdot \|$ magja, azaz, a $\| \cdot \|$: homomorfizmus által indukált kongruencia reláció.

$$\theta = \text{mag}(\| \cdot \|) = \{(v, w) \in M^* \times M^* \mid \|v\| = \|w\|\} .$$

Az $S \subseteq M^* \times M^*$ sztring átíró rendszer az alábbi 13 szabályból áll.

- | | |
|------------------------------------|------------------------------------|
| (1) $(DTA \cdot DT^R, DT^R)$ | (2) $(DT^R \cdot DTA, DT^R)$ |
| (3) $(DT \cdot DT^R, DT^R)$ | (4) $(DT^R \cdot DT, DT^R)$ |
| (5) $(DT^{DR} \cdot DT^R, DT^R)$ | (6) $(DT^R \cdot DT^{DR}, DT^R)$ |
| (7) $(DT^R \cdot DT^R, DT^R)$ | (8) $(DTA \cdot DT^{DR}, DT^{DR})$ |
| (9) $(DT^{DR} \cdot DTA, DT^{DR})$ | (10) $(DT \cdot DT^{DR}, DT^{DR})$ |
| (11) $(DT \cdot DT, DTA \cdot DT)$ | (12) $(DT \cdot DTA, DT)$ |
| (13) $(DTA \cdot DTA, DTA)$ | |

Tekintsük a $\rho : M \rightarrow \{1, 2, \dots\}$ súlyfüggvényt, ahol: $\rho(DTA) = 1$ és $\rho(DT) = \rho(DT^{DR}) = \rho(DT^R) = 2$. Nyilvánvaló hogy S ρ -súlycsökkentő. Így S termináló (noetheri). A [11] cikkben az S sztring átíró rendszerre igazoltuk hogy



1. ábra $\|N\|$ tartalmazási diagramja.

- Minden $(\alpha, \beta) \in S$ szabályra, $\|\alpha\| = \|\beta\|$, azaz, $(\alpha, \beta) \in \theta$.
- Így $\leftrightarrow_S^* \subseteq \theta$.
- $N = IRR(S)$.

Itt $N \subseteq M^*$ az alábbi módon definiált:

$$\begin{aligned}
N = \{ & \lambda, DTA, DT, DTA \cdot DT, DT^R \} \cup \\
& \{ (DT^{DR})^n \mid n \geq 1 \} \cup \\
& \{ (DT^{DR})^n \cdot DT \mid n \geq 1 \} .
\end{aligned}$$

Az 1. ábrán látható az $\|N\| = \{ \|w\| \mid w \in N \}$ halmaznak a tartalmazási diagramja. A diagramról leolvasható hogy $\|N\|$ elemeit egy valódi hierarhiába rendeztük a DT^R osztályon belül.

A fentiekből kaptuk hogy $\leftrightarrow_S^* = \theta$, és $[M] = \|N\|$. Az első egyenlőség azt jelenti hogy S elemei, egyenlőségekként felfogva, generálják az összes $[M]$ feletti érvényes egyenlőséget. A második egyenlőség azt jelenti hogy N elemei $[M]$ elemeit reprezentálják.

A fenti eredményekből következik hogy hogy meg tudunk adni egy olyan lineáris időkomplexitású algoritmust amely tetszőleges $Y_1, \dots, Y_m, Z_1, \dots, Z_n \in M$ fatranszformáció osztályokra eldönti hogy az alábbi négy, egymást kizáró és minden lehetséges esetet kimerítő reláció közül melyik áll fenn.

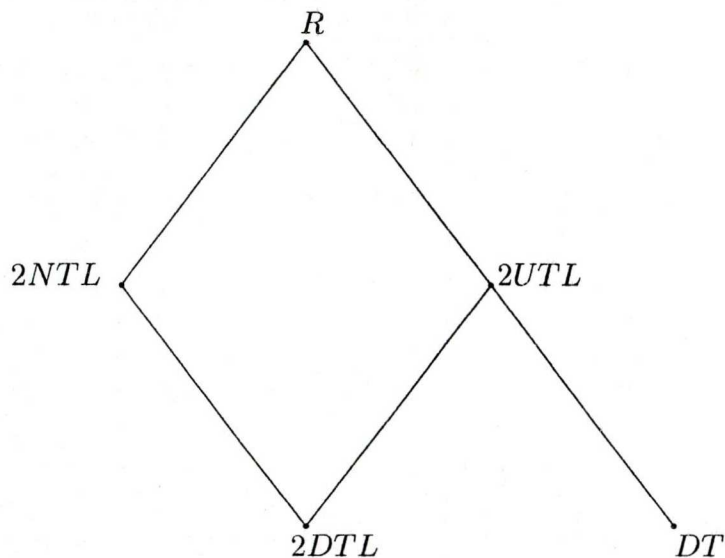
- (i) $Y_1 \circ \dots \circ Y_m = Z_1 \circ \dots \circ Z_n$, (ii) $Y_1 \circ \dots \circ Y_m \subset Z_1 \circ \dots \circ Z_n$,
 (iii) $Z_1 \circ \dots \circ Z_n \subset Y_1 \circ \dots \circ Y_m$, (iv) $Y_1 \circ \dots \circ Y_m \not\subseteq Z_1 \circ \dots \circ Z_n$.

Itt $A \not\subseteq B$ azt jelenti hogy sem $A \subseteq B$ sem $B \subseteq A$ nem teljesül, azaz A és B összehasonlíthatatlanok.

3 Két utas fabejáró szűkítésű top-down fatranszformáció osztályok

A top-down fatranszformátort más típusú reguláris szűkítéssel is el lehet látni. A [14] dolgozatban két utas fabejáró nyelveket választottunk erre a célra. Kamimura és Slutzki [7], [8], [9] bevezett és tanulmányozott három fanyelv osztályt, mind a három valódi részosztálya a felismerhető fanyelvek osztályának. Ez a három osztály a determinisztikus két utas fabejáró automaták által felismert fanyelvek osztálya, amelyet determinisztikus két utas fabejáró fanyelvek osztályának nevezünk és amelyet $2DTL$ jelöl; továbbá a nemdeterminisztikus két utas fabejáró automaták által felismert fanyelvek osztálya, amelyet nemdeterminisztikus két utas fabejáró fanyelvek osztályának nevezünk és amelyet $2NTL$ jelöl; és végül az univerzális két utas fabejáró automaták által felismert fanyelvek osztálya amelyet univerzális két utas fabejáró fanyelvek osztályának nevezünk és amelyet $2UTL$ jelöl. Kamimura és Slutzki [7], [8], [9] összehasonlították az általuk definiált fanyelv osztályokat egymással és az R és DR nyelv osztályokkal, lásd a 2. ábrát.

A [14] cikkben top-down fatranszformátorokat determinisztikus, nemdeterminisztikus és univerzális két utas fabejáró szűkítéssel láttuk el. A megfelelő fatranszformáció osztályokat T^{2DTL} , T^{2NTL} és T^{2UTL} jelöli. A determinisztikus és erősen determinisztikus top-down fatranszformátor osztályok által indukált fatranszformáció osztályokat pedig DT^{2DTL} , DT^{2NTL} és DT^{2UTL} , illetve $D_s T^{2DTL}$, $D_s T^{2NTL}$ és $D_s T^{2UTL}$ jelöli. A [14] cikkben a két utas fabejáró automaták felismerő képességét jellemeztük. Megmutattuk hogy minden $Y \in \{2DTL, 2NTL, 2UTL\}$ fanyelv osztályra az $R = \text{dom}(DT \circ Id(Y))$ egyenlőség teljesül, ahol $Id(Y)$ fatranszformáció osztály elemeit úgy kapjuk meg hogy az identitás relációt megszorítjuk az Y osztály elemeire. Továbbá a Z fatranszformáció osztály elemeinek értelmezési tartományainak az osztályát $\text{dom}(Z)$ jelöli. Ez az eredmény szemléletesen azt jelenti hogy egy A nemdeterminisztikus top-down faautomatát két lépésben lehet szimulálni. Az első lépésben egy determinisztikus top-down fatranszformátor párhuzamos módon kiszámítja \mathcal{A} összes számítását az input fán és a második lépésben egy két utas fabejáró automata \mathcal{A} -nak egy elfogadó kiszámítását keresi. Ennek az eredménynek a felhasználásával a következő dekompozíciós eredményeket kaptuk. Minden $X \in \{ , D \}$ szóra



2. ábra Felismerhető fanyelvek, determinisztikus top-down faautomatákkal felismerhető fanyelvek és a két utas fabejáró fanyelvek osztályának tartalmazási diagramja.

(prefixre) és minden $Y \in \{2DTL, 2NTL, 2UTL\}$ fanyelv osztályra $XT^R = DT \circ XT^Y$ és $FTA \circ DT = DT \circ D_s T^Y$, ahol FTA jelöli a nemdeterminisztikus top-down faautomaták által indukált transzformációk osztályát, $FTA \circ DT$ pedig Rounds [10] transzformációs rendszere.

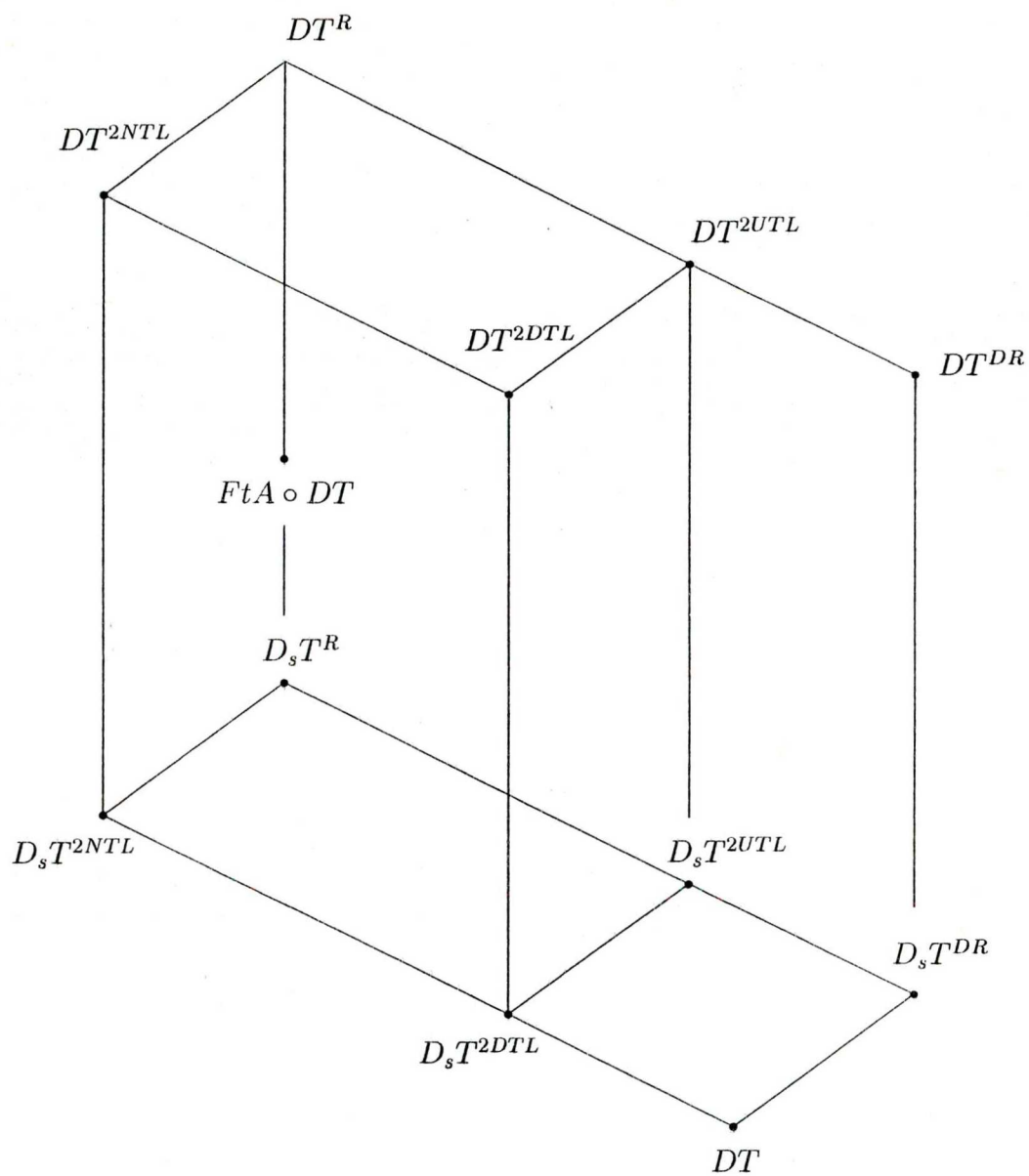
A [14] dolgozatban lerajzoltuk továbbá a különböző szűkítésű determinisztikus és erősen determinisztikus top-down fatranszformáció osztályok tartalmazási diagramját, lásd 3. ábra. A determinisztikus, nemdeterminisztikus és univerzális két utas fabejáró szűkítésű nemdeterminisztikus, determinisztikus és erősen determinisztikus top-down fatranszformáció osztályok kompozícióra való zártságát is megvizsgáltuk. Behatároztuk hogy minden $Y \in \{2DTL, 2NTL, 2UTL\}$ szóra

$$D_s T^Y \subset (D_s T^Y)^2 = (D_s T^Y)^3 = FTA \circ DT,$$

$$DT^Y \subset (DT^Y)^2 = (DT^Y)^3 = DT^R$$

és

$$(T^Y)^k \subset (T^Y)^{k+1}, \quad k \geq 1.$$



3. ábra Két utas fabejáró és reguláris szűkítésű top-down fatranszformáció osztályok tartalmazási diagramja.

Irodalom

- [1] J. Engelfriet, Top-down tree transducers with regular look-ahead, *Mathematical Systems Theory*, **10** (1977) 289-303.
- [2] J. Engelfriet, Three hierarchies of transducers, *Mathematical Systems Theory* **15** (1982) 95-125.
- [3] Z. Fülöp és S. Vágvölgyi, Top-down tree transducers with deterministic top-down look-ahead, *Inform. Process. Lett.* **33** (1989) 3-5.
- [4] Z. Fülöp és S. Vágvölgyi, Variants of top-down tree transducers with look-ahead, *Mathematical Systems Theory* **21** (1989) 125-145.
- [5] Z. Fülöp és S. Vágvölgyi, Iterated deterministic top-down look-ahead, in *Proc. FCT'89, Lecture Notes in Computer Science*, Vol. **380**, Springer-Verlag, 1989, Berlin, pp. 175-184.
- [6] E.T. Irons, A syntax directed compiler for ALGOL 60, *Communications of the ACM* **4** (1961) 51-55.
- [7] T. Kamimura, Tree automata and attribute grammars, *Inform. and Control* **57** (1983) 1-20.
- [8] T. Kamimura és G. Slutzki, Parallel and two-way automata on directed ordered acyclic graphs, *Inform. and Control* **49** (1981) 10-51.
- [9] G. Slutzki, Alternating tree automata, *Theoret. Comput. Sci.* **41** (1985) 305-318.
- [10] W. C. Rounds, Mappings and grammars on trees, *Math. Systems Theory* **4** (1970) 257-287.
- [11] G. Slutzki és S. Vágvölgyi, A Hierarchy of Deterministic Top-down Tree Transformations, *Proc. FCT '93, Lecture Notes in Computer Science*, Vol. **710** (Springer-Verlag, Berlin, 1993) 440-451, a teljes változat a *Mathematical Systems Theory* újságban fog megjelenni.
- [12] G. Slutzki és S. Vágvölgyi, Deterministic Top-down Tree Transducers with Iterated Look-ahead, *Theoret. Comput. Sci.* **143** (1995) 285-308.
- [13] J. W. Thatcher, Generalized² sequential machine maps, *J. Comput. System Sci.* **4** (1970) 339-367.
- [14] S. Vágvölgyi, Top-down tree transducers with two-way tree walking look-ahead, *Theoret. Comput. Sci.* **93** (1992) 43-74.