

Relația limbajelor contextuale cu concatenare cu alte familii de limbaje contextuale

Lector dr. Florin Fortiș
Universitatea de Vest, Timișoara
Asistent Alexandra Fortiș
Universitatea Tibiscus, Timișoara

ABSTRACT. In this paper we are studying the properties for some families of contextual languages with catenation. The study concerns some classical properties (see [13]) for contextual languages and some new properties characteristic to these families of languages. The results in this paper continues the work from [7]. Using these properties we will be able to establish some necessary conditions and pumping properties for families of contextual languages with catenation.

1 Introducere

În definiția originală a gramaticilor contextuale selective, $G = (V, A, C, \varphi)$, limbajul generat este definit ca fiind cel mai mic limbaj satisfăcând următoarele proprietăți:

1. fiecare cuvânt din A se găsește în $L(G)$;
2. pentru fiecare cuvânt $x \in L(G)$ și pentru un cuvânt $z \in L(G)$ și un context $(u, v) \in \varphi(x)$, cuvintele uxv, xz, zx se găsesc de asemenea în $L(G)$.

În acest mod, definiția poate fi utilizată pentru a oferi suportul pentru modelarea unor construcții elaborate în limbajele naturale. Prin introducerea operației de concatenare, alături de operația de utilizare a contextelor, pot fi obținute construcții mai complexe, urmând modelul de formare a propozițiilor, frazelor sau textelor. Noile dispozitive generative, gramatici contextuale cu concatenare, sunt utilizate pentru a extinde puterea

generativă a dispozitivelor clasice, gramatici contextuale (Marcus) (vezi [For00a], [For01b], [For01c], [Mar69a], [Pău82]).

Această lucrare realizează o investigare a relației gramaticilor contextuale cu concatenare în raport cu alte familii de gramatici contextuale completând rezultatele deja prezentate în lucrarea [For01b].

2 Ierarhia limbajelor contextuale cu concatenare

Vom nota în cele ce urmează cu $CSEL_c$, respectiv $CSEL_{in,c}$ familiile limbajelor contextuale cu concatenare a căror funcție de selecție este calculabilă. Direct din definițiile diferitelor clase de limbaje contextuale cu concatenare rezultă următoarea

Lema 2.1. *Au loc următoarele relații*

$$\begin{aligned} CCON &\subseteq CSEL_c \subseteq CSEL \\ CCON_{in} &\subseteq CSEL_{in,c} \subseteq CSEL_{in} \end{aligned}$$

În plus, toate incluziunile sunt stricte.

Faptul că aceste incluziuni sunt stricte rezultă din exemplele considerate. Astfel, limbajul $a^+ \cup b^+$ nu este $CCON$ sau $CCON_{in}$ însă este $CSEL_c$ și $CSEL_{in,c}$. De asemenea, printr-o justificare similară cu cea din [Păun 97,

Lema 5.11] putem pune în evidență faptul că incluziunile $CSEL_c \subseteq CSEL$ și $CSEL_{in,c} \subseteq CSEL_{in}$ sunt stricte.

Mai departe prezentăm un rezultat care separă complet cele două clase de gramatici, după tipul derivării (internă sau externă)

Lema 2.2. *Fiecare dintre familiile $CCON$, $CSEL_c$, $CSEL$ este incomparabilă cu fiecare dintre familiile $CCON_{in}$, $CSEL_{in,c}$, $CSEL_{in}$.*

Demonstrație. Pentru început vom considera limbajele

$$L_1 = a * ba * ba *$$

și

$$L_2 = \{xmi(x) \mid x \in \{a, b\}^*\}.$$

Atunci $L_1 \in CSEL_{in} \setminus CSEL$ iar $L_2 \in CSEL \setminus CSEL_{in}$.

Mai departe fie limbajul

$$L_3 = \{ba, ba^2, \dots, ba^k\}, k \geq 2$$

Acest limbaj este $CCON$ însă nu și $CCON_{in}$.

Pentru relația inversă, considerăm limbajul generat de următoarea gramatică

$$G_4 = (\{a, b, c\}, \{abc\}, \{(\lambda, \lambda)\})$$

Pentru un cuvânt $w \in L(G_4)$ are loc relația $N_a(w) = N_b(w) = N_c(w)$. De asemenea, cuvântul $x = a^n (bc)^{n-1} b (ab)^{n-1} c^n, n \geq 1$ se găsește în $L(G_4)$. Pe de altă parte, nu există $x_1, x_2 \in L(G_4)$ astfel încât $x = x_1 x_2$. Mai mult, acest limbaj nu posedă proprietatea EBSC, așadar limbajul $L(G_4)$ nu poate fi CCON.

3 Relația cu familii de limbaje contextuale

3.1 Relația cu familiile EC, ECC

Pentru a pune în evidență relația dintre familiile de limbaje contextuale cu concatenare și limbajele familiilor EC, ECC, să observăm pentru început că, în timp ce limbajele EC sunt 1-liniare, limbajele CCON sunt independente de context, așa cum rezultă din construcția de mai jos. Mai mult, familia CCON conține limbaje de tip Dyck.

Lema 3.1. *Orice limbaj CCON este CF.*

Demonstrație. Fie L un limbaj contextual cu concatenare generat de gramatica $G = (V, A, C)$. Definim gramatica independentă de context $G_1 = (\{S\}, V, S, P)$ unde P conține producțiile

$$\begin{aligned} S &\rightarrow SS \\ S &\rightarrow x, \quad x \in V \\ S &\rightarrow uSv, \quad (u, v) \in C \end{aligned}$$

Relația $L(G) = L(G_1)$ poate fi verificată imediat.

Să considerăm mai departe un limbaj EC simplu, $L_1 = \{a^n b^n \mid n \geq 1\}$. Acest limbaj este EC, dar nu este și CCON, așa cum rezultă din demonstrația din următoarea

Lema 3.2. *Familiile CCON și EC sunt incomparabile.*

Demonstrație. Într-adevăr, pentru ca limbajul L_1 de mai sus să fie CCON, pentru $a^n b^n, a^m b^m \in L_1$, cu $m, n \geq 1$ ar urma că $a^n b^n a^m b^m \in L_1$, cu $m, n \geq 1$.

Considerăm acum gramatica contextuală cu concatenare

$$G_2 = (\{a, b\}, \{\lambda\}, \{(a, b), (b, a)\})$$

Pentru un cuvânt $x \in L(G_2)$ are loc relația $N_a(x) = N_b(x)$. Mai departe, deoarece $a^n b^{n+m} a^m \in L(G_2)$, pentru $m, n \geq 1$, suficient de mari, urmează că limbajul $L(G_2)$ nu posedă proprietatea *EBS*, deci acest limbaj nu poate fi *EC* (dar nici măcar *ECC*).

Ținând seama de definiția limbajelor *CSEL* și de rezultatele din secțiunea anterioară nu este greu de stabilit acum

Lema 3.3. *Orice limbaj EC este CSEL. Mai mult, orice limbaj ECC este CSEL.*

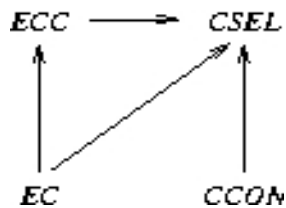
Demonstrație. Este suficient să observăm că simpla blocare a operației de concatenare face ca limbajul generat să devină *ECC*. Prin urmare $EC, ECC \subseteq CSEL$.

Idea "blocării" concatenărilor de cuvinte prin intermediul selectorului (λ, λ) este utilizată în mai multe locuri în această lucrare. Blocarea concatenărilor se realizează efectiv prin eliminarea contextelor (λ, λ) din definiția gramaticii. Astfel, gramatica este "retrogradată" la gramatica contextuală (simplă, selectivă, interioară) pe care se bazează în realitate întreaga construcție a gramaticii contextuale cu concatenare. De asemenea, această operație de blocare a concatenărilor pune în evidență importanța contextului (λ, λ) în generarea cuvintelor.

Mai departe, ținând seama de definițiile familiilor de limbaje $CCON_{in}$ și $CSEL_{in}$, nu este dificil de stabilit următorul rezultat:

Lema 3.4. *Familiile $CCON_{in}$ și $CSEL_{in}$ nu sunt comparabile cu familia ECC (dar nici cu familia EC).*

Rezultatele din această secțiune pot fi acum sintetizate în următoarea diagramă



3.2 Relația cu familiile *IC*, *ICC*

Așa cum este de așteptat, nici legătura cu familiile de limbaje interioare nu

se dovedește a fi mult mai puternică,

Lema 3.5. *Famiiliile CCON și IC sunt incomparabile.*

Demonstrație. Pentru început, să reamintim că au loc relațiile $FIN \subseteq IC, ICC$.

Mai departe, considerăm gramatica contextuală cu concatenare

$$G_1 = (\{a, b\}, \{ab\}, \{(\lambda, \lambda)\})$$

Limbajul generat de această gramatică este $L(G_1) = \{(ab)^n \mid n \geq 1\}$. Să observăm că pentru orice cuvânt $x \in L(G_1)$, $N_a(x) = N_b(x)$, prin urmare orice context al unei gramatici IC , fie aceasta G_2 , care generează acest limbaj va avea această proprietate. Mai mult, în orice subcuvânt al unui cuvânt $x \in L(G_1)$ fiecare apariție a simbolului a se află înaintea oricărei apariții a unui simbol b utilizat. Prin urmare, pentru o valoare a lui n suficient de mare, gramatica G_2 poate genera cuvinte din afara limbajului $L(G_1)$.

Limbajul din demonstrația de mai sus este ICC . Totuși, între familiile $CCON$ și ICC are loc următoarea

Lema 3.6. *Famiiliile CCON și ICC sunt incomparabile.*

Demonstrație. Considerăm de această dată următoarea gramatică contextuală cu concatenare

$$G_3 = (\{a, b, c\}, \{ab, ab^3, a^3b\}, \{(c, c)\})$$

Pentru ca o gramatică contextuală interioară cu selecție să genereze limbajul $L(G_3)$, aceasta trebuie să fie capabilă să construiască cuvinte de forma $c(ab)^n c$ cu $n \geq 1$. În același timp însă în $L(G_3)$ urmează să se găsească și cuvinte de forma $a^3b(ab)^{n-2}ab^3, (ab)^n$.

O gramatică contextuală interioară capabilă să genereze cuvintele $c(ab)^n c$ va genera, inevitabil, și cuvinte de forma $a^2c(ab)^n cb^2$, cuvinte care nu se află în $L(G_3)$.

Același tip de rezultat are loc și pentru relația dintre familia $CSEL$ și familiile ICC, IC .

Lema 3.7. *Familia CSEL este incomparabilă cu fiecare dintre familiile IC și ICC.*

Demonstrație. Considerăm limbajul

$$L = \{x \in \{a, b, c\}^* \mid N_a(x) = N_b(x) = N_c(x)\}$$

Acest limbaj este interior (selectiv) însă nu este și contextual selectiv cu

concatenare. Pentru ca limbajul L să fie contextual selectiv cu concatenare ar trebui să putem genera cuvinte de forma $a^n b^n c^n$ în condițiile în care $N_a(x) = N_b(x) = N_c(x)$ pentru orice cuvânt $x \in L$.

Ținând acum seama de relațiile existente între $CCON$ și IC , ICC , obținem rezultatul enunțat.

Situația se schimbă sensibil în momentul în care avem de-a face cu gramatici contextuale cu concatenare cu derivare internă. Mai exact, are loc

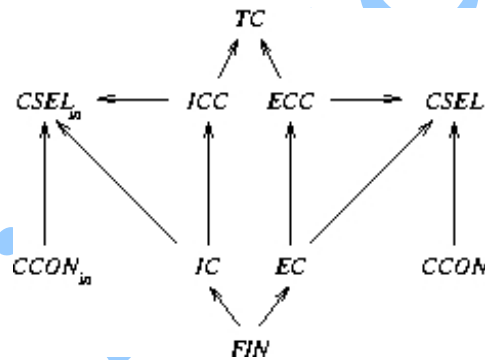
Lema 3.8. *Între familiile IC , ICC și $CSEL_{in}$ are loc relația*

$$IC, ICC \subseteq CSEL_{in}$$

Mai mult, relația de incluziune este strictă.

Demonstrație. Din nou, va fi suficient să eliminăm posibilitatea de a realiza concatenări, prin suprimarea contextelor de forma (λ, λ) , pentru a obține incluziunile $IC, ICC \subseteq CSEL_{in}$.

Rezultatele din această secțiune, împreună cu rezultatele din secțiunea anterioară, pot fi acum sintetizate în următoarea diagramă



3.3 Relația cu familia TC

Mai departe vom trece la stabilirea relațiilor dintre familiile de limbaje contextuale cu concatenare și familia limbajelor TC .

Pentru început, să ne aducem aminte că familia $CCON$ posedă proprietatea IBS, prin urmare are loc

Lema 3.9. *Familia $CCON$ este (strict) inclusă în familia TC .*

Un rezultat similar poate fi stabilit și pentru familia $CCON_{in}$. Rezultatul pentru acest caz se bazează din nou pe proprietatea IBS.

În ceea ce privește familia $CSEL$ are loc

- Workshop on Symbolic and Numerical Algorithms for Scientific Computation, SYNASC'2000, Timișoara, 2000
- [For01a] **F. Fortiș**, *FBICC(FIN) in the Chomsky Hierarchy*, Ann. Univ. Timișoara, 39, 1, 2001
- [For01b] **F. Fortiș**, *Contextual Grammars with Catenation. Generative Capacity*, Ann. Univ. Timișoara, 39, 2, 2001
- [For01c] **F. Fortiș**, *Results of Decidability for Contextual Grammars with Catenation*, Ann. Univ. Timișoara, 39, 2, 2001
- [For01d] **F. Fortiș**, *On the generative capacity of contextual grammars with catenation*, Informatica, Lithuania, 10, 2, 2001
- [Mar69] **S. Marcus**, *Contextual grammars*, Rev. Roum. Math. Pures Appl., 14, 1969
- [Mar69] **S. Marcus**, *Deux types nouveaux de grammaires generatives*, Cah. Ling. Th. Appl., 6, 1969
- [MVP98] **C. Martin-Vide, Gh. Păun**, *Structured contextual grammars*, Grammars, 1, 1, 1998
- [Pău79a] **Gh. Păun**, *Marcus' contextual grammars and languages. A survey*, Rev. Roum. Math. Pures Appl., 24, 1979
- [Pău79b] **Gh. Păun**, *On a prolongation operation of languages*, Bull. Math. Soc. Sci. Math. Roumanie, 23(71), 1979
- [Pău82] **Gh. Păun**, *Contextual Grammars*, Editura Academiei, 1982