Polinomok értékkészlete minden számrendszerben nem környezet–független, vagy reguláris

Pontos bibliográfiai adatok:
Horváth Sándor. Ranges of polinomials are not context-free in any number system. To appear in PuMa, June 1998.

Rövid cím: Polinomok értékkészletéről.

Tekintsünk egy egészeken értelmezett p polinomot és egy $B \geq 1$ számot! Definiáljunk egy L nyelvet a következőképp: a nyelv szavai a p polinom értékkészletének nemnegatív egész elemei B alapú számrendszerben felírva. A cikk azzal foglalkozik, hogy az ily módon kapott L nyelvek mely Chomsky osztályokba tartozhatnak. Az eredmény az, hogy (a polinomra tett bizonyos megkötések mellett) elsőfokú polinom esetén a nyelv reguláris, magasabb fokszám esetén pedig nem környezet–független, de környezetfüggő, függetlenül a számrendszer alapjától és a polinomtól.

Formálisan, ha p egy polinom, és $B \geq 1$ egy szám, akkor jelölje $L(p, B) = (p(\mathcal{N}) \cap \mathcal{N})_B$, ahol az alsó index B a B alapú számrendszerben való felírást jelenti, $\mathcal{N} = \{0, 1, 2, \ldots\}$.

Tétel. *Legyen $p(x) = a_r x^r + \ldots + a_1 x + a_0$ egy tetszőleges egész–együtthatós polinom, melyre $r \geq 1$ és $a_r > 0$, továbbá legyen $B \geq 1$ egész! Ekkor $L(p, B)$ nem környezet–független, ha $r \geq 2$, és reguláris, ha $r = 1$.*

A tétel bizonyításában a cikk elsősorban a Bar–Hillel lemmára, illetve véges automaták elemi tulajdonságaira hivatkozik. A bizonyítás szétválik az $r \geq 3$, $r = 2$ és $r = 1$ esetekre, miközben felhasznál két egyszerű lemmát.

A cikk második részében az előző tétel erősebb változata szerepel: Belátjuk, hogy a polinomra tett feltételek gyengítése mellett is teljesül az állítás. Nevezetesen a polinom lehet tetszőleges komplex együtthatós, amennyiben $p(n) \in \mathcal{N}$ teljesül véges sok kivételével minden $n \in \mathcal{N}$ esetén.

A dolgozat egy sejtéssel zárul, melyben a polinomra tett megszorítás még gyengébb, és az állítás továbbra is az, hogy $L(p, B)$ sohasem lehet nem reguláris környezet–független.

Egy „kicserélési lemma” környezet-független nyelvekre

Pontos bibliográfiai adatok:
William Ogden, Rockford J. Ross, and Karl Winklmann. An „interchange lemma” for context-free languages. *SIAM Journal on Computing*, 14(2):410–415, May 1985.

Rövid cím: Kicserélési lemma.

A cikk egy új szükséges feltételt ad környezet–független nyelvekre, majd példát is ad az alkalmazására: a lemma segítségével belátja, hogy nem környezet–független egy olyan nyelvről, amely sokáig ellenállt minden hasonló bizonyítási kísérletnek.

Tétel. *Legyen L egy T abc-jű környezet–független nyelv. Ekkor létezik olyan $c > 0$ konstans, hogy minden $n \geq 2$ egészre, minden $Q \subseteq L \cap T^n$ halmazra és minden olyan m egészre, melyre $n \geq m \geq 2$ létezik $k \geq |Q|/(cn^2)$ egész és $z_1, \ldots, z_k \in Q$ szavak a következő tulajdonságokkal:*

1. $z_i = w_i x_i y_i$, $i = 1, \ldots, k$,
2. $|w_1| = |w_2| = \cdots = |w_k|$,
3. $|y_1| = |y_2| = \cdots = |y_k|$,
4. $m \geq |x_1| = |x_2| = \cdots = |x_k| > m/2$, *és*
5. $w_i x_j y_i \in L$ *minden* $i, j \in \{1, 2, \ldots, k\}$ *esetén.*

A bizonyítás Chomsky normálformában levő nyelvtanokhoz tartozó levezetési fák levelei számlálásának segítségével történik.

A cikk az alkalmazásra is példát ad: megmutatja, hogy a „dadogó” szavak (azaz az $xyyz$ alakú szavak, ahol y nem üres) nyelve hárombetűs abc felett nem környezet–független.

Primitív szavak nyelve

Pontos bibliográfiai adatok:
P. Dömösi, S. Horváth, M. Ito, L. Kászonyi, and M. Katsura. Formal languages consisting of primitive words. *Lecture Notes in Computer Science*, 710:194–, 1993.

Rövid cím: Primitív szavak nyelve.
Megjegyzés: A cikket ketten is előadhatják.

Jelöljük Q-val a primitív szavak nyelvét, azaz azokat a szavakat, amelyek nem állnak elő más szó hatványaként. (Formálisan $Q = T^+ \setminus \{w^m \mid w \in T^*, m \geq 2\}$, ahol T egy legalább kétbetűs abc.) Sokat kutatott téma, hogy vajon Q környezet–független-e, a sejtés az, hogy nem. Jelen cikk nem dönti el a kérdést, de megmutatja, hogy Q teljesít két erős feltételt, mindkettő szükséges feltétele a környezet–független nyelveknek.

Az első feltétel Bader és Moura feltételének erősítése (erős BM feltétel). Minden környezet–független nyelv rendelkezik a következő tulajdonsággal:

Definíció. *Legyen $L \subseteq T^*$ egy nyelv. Akkor mondjuk, hogy L teljesíti az* erős BM feltételt, *ha létezik olyan $n \geq 2$ egész, hogy minden $z \in L$ szóra, ha z–ben d pozíció* kitüntetett *és $e \geq 0$ pozíció* kizárt, *ahol $d > n^{1+e}$, akkor z faktorizálható a következő módon: $z = uvwxy$, ahol*

1. *w tartalmaz kitüntetett pozíciót,*
2. *vagy u és v, vagy x és y mindegyike tartalmaz kitüntetett pozíciót,*
3. *vx nem tartalmaz kizárt pozíciót,*
4. *ha a kitüntetett és kizárt pozíciók száma vwx-ben rendre r és s, akkor $r \leq n^{1+s}$, végül*
5. *minden $m \geq 0$ számra $uv^m wx^m y \in L$.*

Figyeljük meg, hogy $d = |z|$, $e = 0$ választással a jól ismert Bar-Hillel lemmát kapjuk. A cikk szerzői belátják, hogy Q teljesíti az erős BM feltételt.

A cikk további részében a szerzők megmutatják, hogy Q egy speciális részhalmaza nem környezet–független, de megadnak végtelen sok olyan reguláris R nyelvet, melyekre $Q \cap R$ környezet–független.

A primitív szavak nyelve erősen kicserélhető, de nem lineáris

Pontos bibliográfiai adatok:
Horváth Sándor. Primitive words are strongly interchangeable but nonlinear. Technical Report FBI-HH-B-183/96, Universität Hamburg Fachbereich Informatik, Vogt-Kölnn-Strasse 30, D-22527 Hamburg, Germany, January 1996.

Rövid cím: Erős kicserélhetőség.

Jelöljük Q-val a primitív szavak nyelvét, azaz azokat a szavakat, amelyek nem állnak elő más szó hatványaként. (Formálisan $Q = T^+ \setminus \{w^m \mid w \in T^*, m \geq 2\}$, ahol T egy legalább kétbetűs abc.) Sokat kutatott téma, hogy vajon Q környezet–független-e, a sejtés az, hogy nem. Jelen cikk nem dönti el a kérdést, de négy eredményt közül, ami közelebb hozhatja a megoldást.

A szerző bevezeti az Ogden, Ross és Winklmann kicserélési tulajdonságának egy erősített változatát, majd megmutatja hogy Q rendelkezik ezzel a tulajdonsággal.

Definíció. *Azt mondjuk, hogy egy $L \subseteq T^*$ nyelv* erősen kicserélhető, *ha létezik olyan $c > 0$ konstans, melyre minden $n \geq 2$, $i \geq 0$ és $j \geq 1$ esetén, ahol $j < n$ és $i + j \leq n$, és minden $H \subseteq L \cap T^n$ nem üres halmazra létezik $H' \subseteq H$ a következő tulajdonságokkal:*

1. $|H'| > |H|/c$
2. *$x, y \in H'$ esetén, ha $x = x_1x_2x_3$, $y = y_1y_2y_3$, ahol $|x_1| = |y_1| = i$, $|x_2| = |y_2| = j$, teljesül a következő: $x_1y_2x_3, y_1x_2y_3 \in L$.*

A dolgozat további részében a szerző bizonyítja, hogy a legfeljebb n hosszú primitív szavak sűrűsége exponenciális gyorsasággal tart 0-hoz n növekedtével, valamint azt, hogy Q nem korlátos nyelv, és nem is lineáris.

Szavak kombinatorikus tulajdonságai és a Chomsky osztályok kapcsolata

Pontos bibliográfiai adatok:
Dömösi Pál et al. Some combinatorial properties of words, and the Chomsky hierarchy. In *Proc. Internat. Conf. „Words, Languages and Combinatorics, II”, Kyoto, Japan; eds.: M. Ito and H. Jürgensen*, pages 105–123. Singapore: World Scientific, 1992.

Rövid cím: Kombinatorikus tulajdonságok.
Megjegyzés: A cikket ketten is előadhatják.

Jelöljük Q-val a primitív szavak nyelvét, azaz azokat a szavakat, amelyek nem állnak elő más szó hatványaként. (Formálisan $Q = T^+ \setminus \{w^m \mid w \in T^*, m \geq 2\}$, ahol T egy legalább kétbetűs abc.) Sokat kutatott téma, hogy vajon Q környezet–független-e, a sejtés az, hogy nem. Jelen cikk nem dönti el a kérdést, de belátja, hogy Q teljesít három olyan feltételt, az Ogden, az erős Bader–Moura és a Sokolowsky feltételt, amelyet minden környezet–független nyelv teljesít. A cikk egy negyedik feltételt is vizsgál, ami szükséges feltétele minden környezet–független nyelvnek, a Parikh szemilinearitást. Megmutatja, hogy $Q \cap (a^*b)^n$ (ahol $n \geq 1$ egész) Parikh szemilineáris, tehát ez a támadási kísérlet is kudarcot vall. A dolgozat utolsó része általánosan vizsgál bizonyos nyelvi tulajdonságokat, különös tekintettel azok eldönthetőségére és elhelyezkedésükre a Kleene féle aritmetikai hierarchiában.

Definíció. *Azt mondjuk, hogy egy $L \subseteq T^*$ nyelv teljesíti a* Ogden feltételt, *ha létezik olyan $n \geq 2$ egész, melyre $z \in L$ esetén, ha legalább $n+1$ pozíciót z-ben kitüntetettnek tekintünk, akkor z felbontható $z = uvwxy$ alakban, ahol teljesülnek a következők:*

1. *vagy u, v és w, vagy w, x és y mindegyike tartalmaz kitüntetett pozíciót,*
2. *vwx legfeljebb n kitüntetett pozíciót tartalmaz,*
3. *minden $m \geq 0$ számra $uv^mwx^my \in L$.*

Tétel. *Minden környezet–független nyelv teljesíti az Ogden feltételt.*

Definíció. *Azt mondjuk, hogy egy $L \subseteq T^*$ nyelv teljesíti a* Sokolowsky feltételt, *ha minden $Y \subseteq T^*$, $|Y| \geq 2$, és minden $u_1, u_2, u_3 \in T^*$ esetén, ha $\{u_1xu_2xu_3 \mid x \in Y^+\} \subseteq L$, akkor létezik olyan $x', x'' \in Y^+$, $x' \neq x''$, melyre $u_1x'u_2x''u_3 \in L$.*

Tétel. *Minden környezet–független nyelv teljesíti a Sokolowsky feltételt.*

Környezet–független nyelvekre adott kicserélési és pumpáló lemmák

Pontos bibliográfiai adatok:
R. Boonyavatana and G. Slutzki. The interchange or pump (DI)lemmas for context-free languages. *Theoretical Computer Science*, 56(3):321–338, March 1988.

Rövid cím: Feltételek összehasonlítása.
Megjegyzés: A cikket ketten is előadhatják.

A cikk összehasonlítja a kicserélési lemmát különböző pumpáló lemmákkal: a Bar–Hillel félével, Ogdenével, ennek általánosított változatával, a fentiek lineáris változatával, valamint a Sokolowsky féle feltételekkel. Ezen felül definiál egy kicserélési feltételt lineáris környezet–független nyelvekre, és összehasonlítja a többi feltétellel. Az eredmények megmutatják, hogy a kicserélési feltételek nem összehasonlíthatóak a pumpáló feltételekkel és a kiterjesztett Sokolowsky feltétellel, de határozottan erősebbek a Sokolowsky feltételnél.

Definíció. *Azt mondjuk, hogy egy $L \subseteq T^*$ nyelv teljesíti a* lineáris kicserélési feltételt, *ha létezik olyan $c > 0$ konstans, hogy minden $n \geq m \geq 2$ egészekre, és minden $Q \subseteq L \cap T^n$ halmazra létezik $R \subseteq Q$ a következő tulajdonságokkal:*

1. $|R| \geq |Q|/(cn)$, *és*
2. *valamilyen $i \geq 0$ számra $\{xwy \mid \exists u, v, z \in T^* : |x| = |u| = i, |w| = |z| = m, xzy, uwv \in R\} \subseteq L \cap T^n$.*

Definíció. *Azt mondjuk, hogy egy $R \subseteq T^* \times T^*$ reláció* nem korlátos, *ha minden m számra léteznek olyan $x, y \in T^*$ szavak, melyre $|x|, |y| \geq m$, és $R(x, y)$ fennáll.*

Definíció. *$x, y \in T^*$ esetén $x < y$ azt jelenti, hogy x-et úgy kaphatjuk y-ból, hogy néhány betűjét kitöröljük. Ha a törölt szimbólumok y utolsó (első) m szimbóluma közül kerülnek ki, azt így jelöljük: $x <^m y$ ($x\ {}^m\!< y$). Szavak párjaira definiáljuk a következő relációt:*

$$(x', y') <_m (x, y) \Leftrightarrow (x' < x \wedge y' = y) \vee (x' = x \wedge y' < y) \vee (x' <^m x \wedge y'\ {}^m\!< y).$$

Definíció. *Azt mondjuk, hogy egy $L \subseteq T^*$ nyelv teljesíti a* kiterjesztett Sokolowsky feltételt, *ha minden $u_1, u_2, u_3 \in T^*$ szóra és minden $R \subseteq T^* \times T^*$ nem korlátos relációra, ha $\{u_1 x u_2 y u_3 \mid R(x, y)\} \subseteq L$, akkor létezik olyan m egész, hogy minden $x, y \in T^*$ esetén*

$$|x|, |y| > m \ \wedge \ R(x, y) \implies \exists x', y' : (x', y') <_m (x, y) \ \wedge \ u_1 x' u_2 y' u_3 \in L.$$

A rekurzívan felsorolható nyelvek jellemzése beszúró–törlő rendszerek segítségével

Pontos bibliográfiai adatok:
Lila Kari, Gheorghe Păun, Gabriel Thierrin, and Sheng Yu. At the crossroads of dna computing and formal languages: Characterizing recursively enumerable languages using insertion-deletion systems. In *Proceedings of the 3rd DIMACS Workshop on DNA Based Computers, held at the University of Pennsylvania, June 23 – 25, 1997*, pages 318–333, June 1997.

Rövid cím: Beszúró–törlő rendszerek.
Megjegyzés: A cikket ketten is előadhatják.

A formális nyelvek elméletében a leggyakoribb generáló mechanizmus az újraírás. A természetes nyelvek vizsgálata során merült fel a gondolat, hogy az újraírás mellett, vagy helyett szövegrészek beszúrásával, vagy törlésével képezzünk új szövegeket. E módszerek a molekuláris számítások szempontjából is rendkívül fontosak, mert különböző enzimek segítségével DNS, illetve RNS molekulákon megvalósíthatóak bizonyos (a környezettől is függő) beszúrások és törlések. A cikk a beszúró–törlő rendszerek több alosztályáról belátja, hogy az általuk generált nyelv megegyezik a rekurzívan felsorolható nyelvekkel.

Definíció. *Beszúró–törlő rendszeren a következő konstrukciót értjük:*

$$\gamma = (V, T, A, I, D),$$

ahol V egy abc, $T \subseteq V$, $A \subseteq V^*$, $I, D \subseteq V^* \times V^* \times V^*$ véges halmazok. T a rendszerünk terminális abc-je, A az axiómák, I a beszúró, D a törlő szabályok halmaza. $(u, z, v) \in I$ azt jelenti, hogy ha egy szóban előfordul az uv részszó, akkor u és v közé beszúrhatjuk a z szót. $(u, z, v) \in D$ azt jelenti, hogy ha egy szóban előfordul az uzv részszó, akkor u és v közül kitörölhetjük a z szót.

$x, y \in V^*$ esetén $x \Longrightarrow y$ pontosan akkor áll fenn, ha a következők egyike teljesül:

- $x = x_1uvx_2$, $y = x_1uzvx_2$, $x_1, x_2 \in V^*$ és $(u, z, v) \in I$
- $x = x_1uzvx_2$, $y = x_1uvx_2$, $x_1, x_2 \in V^*$ és $(u, z, v) \in D$

Jelölje $\Longrightarrow^*$ a $\Longrightarrow$ reláció reflexív, tranzitív lezártját. A γ által generált nyelv:

$$L(\gamma) = \{w \in T^* \mid x \Longrightarrow^* w, x \in A\}.$$

A reverzibilis két–számlálójú gép univerzalitásáról

Pontos bibliográfiai adatok:
Kenichi Morita. Universality of a reversible two-counter machine. *Theoretical Computer Science*, 168(2):303–320, 20 November 1996.

Rövid cím: Két–számlálójú gép.
Megjegyzés: A cikket ketten is előadhatják.

A k–számlálójú gép ($CM(k)$) egy olyan automata, aminek k számláló áll rendelkezésre segédmemóriaként. Minsky megmutatta, hogy bármely Turing gépet lehet szimulálni egy $CM(2)$-beli géppel. Ez a cikk a reverzibilis (visszafelé determinisztikus) számláló gépekkel foglalkozik. Először megmutatja, hogy bármely $CM(k)$ gép szimulálható egy reverzibilis $CM(k+2)$ géppel, majd bebizonyítja, hogy minden reverzibilis $CM(k)$ gép szimulálható egy reverzibilis $CM(2)$-belivel.

Definíció. *k–számlálójú gép alatt a következőt értjük:*

$$M = (k, Q, \delta, q_0, q_f),$$

ahol k a szalagok (vagy számlálók) száma, Q az állapotok véges, nem üres halmaza, $q_0 \in Q$ a kezdőállapot, $q_f \in Q$ a végső állapot. A szalagok abc-je $\{Z, P\}$, ahol P az üres szimbólum. δ a konfiguráció–átmeneti reláció: $\delta \subseteq (Q \times \{1, \dots, k\} \times \{Z, P\} \times Q) \cup (Q \times \{1, \dots, k\} \times \{-, 0, +\} \times Q)$. A $-$, 0, $+$ szimbólumok balra mozgást, helyben maradást, illetve jobbra mozgást jelentenek. A szalagok egy irányba (jobbra) végtelen hosszúak, a legbaloldali pozícióban a Z szimbólum szerepel, az összes többiben pedig a P. (Z a zérus, P a pozitív rövidítése.)

δ minden eleme egy négyes, alakja vagy (q, i, s, q'), vagy (q, i, d, q'), ahol $q, q' \in Q$, $s \in \{Z, P\}$, $d \in \{-, 0, +\}$. A (q, i, s, q') négyes azt jelenti, hogy ha a gép a q állapotban van, és az i. olvasófej alatt az s szimbólum található, akkor átmehet a q' állapotba. A (q, i, d, q') négyes azt jelenti, hogy ha a gép a q állapotban van, átmehet a q' állapotba, és az i. olvasófej balra, jobbra, vagy nem mozog egy pozíciót d értéke szerint.

Nyelvek generálása kétfajta zárójelpár segítségével

Pontos bibliográfiai adatok:
Viliam Geffert. How to generate languages using only two pairs of parentheses. *EIK: Journal of Information Processing and Cybernetics EIK (formerly Elektronische Informationsverarbeitung und Kybernetik)*, 27, 1991.

Rövid cím: Geffert normálforma.
Megjegyzés: A cikket ketten is előadhatják.

A szerző két jól használható normálformát ad a 0. típusú nyelvtanokra.

Tétel. *Minden G nulladik típusú nyelvtanhoz megadható egy vele ekvivalens G' nyelvtan, melynek idő– és térbonyolultsága aszimptotikusan egyenlő a G nyelvtanéval, mindössze* 5 *nemterminális szimbólumot használ (ezek: $S', (,), [,]$), szabályai kettő kivételével környezet–függetlenek. (A két kivétel: $() \to \epsilon$, és $[] \to \epsilon$.)*

Tétel. *Minden G nulladik típusú nyelvtanhoz megadható egy vele ekvivalens G' nyelvtan, melynek idő– és térbonyolultsága aszimptotikusan egyenlő a G nyelvtanéval, mindössze* 5 *nemterminális szimbólumot használ (ezek: S', X, A, B, C), szabályai egy kivételével környezet–függetlenek. (A kivétel: $ABC \to \epsilon$.)*

Formális nyelvek szeminárium, 2000
Tervezett program

A [3], [5], [6], [7], [8], [9] cikkeket ketten is előadhatják.

Szükséges feltételek környezet–független nyelvekre

[1] Horváth Sándor. Ranges of polinomials are not context-free in any number system. To appear in PuMa, June 1998.

[2] William Ogden, Rockford J. Ross, and Karl Winklmann. An "interchange lemma" for context-free languages. *SIAM Journal on Computing*, 14(2):410–415, May 1985.

[3] P. Dömösi, S. Horváth, M. Ito, L. Kászonyi, and M. Katsura. Formal languages consisting of primitive words. *Lecture Notes in Computer Science*, 710:194–, 1993.

[4] Horváth Sándor. Primitive words are strongly interchangeable but nonlinear. Technical Report FBI-HH-B-183/96, Universität Hamburg Fachbereich Informatik, Vogt-Kölnn-Strasse 30, D-22527 Hamburg, Germany, January 1996.

[5] Dömösi Pál et al. Some combinatorial properties of words, and the Chomsky hierarchy. In *Proc. Internat. Conf. "Words, Languages and Combinatorics, II", Kyoto, Japan; eds.: M. Ito and H. Jürgensen*, pages 105–123. Singapore: World Scientific, 1992.

[6] R. Boonyavatana and G. Slutzki. The interchange or pump (DI)lemmas for context-free languages. *Theoretical Computer Science*, 56(3):321–338, March 1988.

Egyéb témák

[7] Lila Kari, Gheorghe Păun, Gabriel Thierrin, and Sheng Yu. At the crossroads of dna computing and formal languages: Characterizing recursively enumerable languages using insertion-deletion systems. In *Proceedings of the 3rd DIMACS Workshop on DNA Based Computers, held at the University of Pennsylvania, June 23 – 25, 1997*, pages 318–333, June 1997.

[8] Viliam Geffert. How to generate languages using only two pairs of parentheses. *EIK: Journal of Information Processing and Cybernetics EIK (formerly Elektronische Informationsverarbeitung und Kybernetik)*, 27, 1991.

[9] Kenichi Morita. Universality of a reversible two-counter machine. *Theoretical Computer Science*, 168(2):303–320, 20 November 1996.

Tartalék cikkek

[10] Arto Salomaa. *Abstact Families of Languages, In: Formal Languages*, chapter IV, pages 123–141. Academic Press, New York, 1973.

[11] Arto Salomaa. *Finite non–deterministic and probabilistic automata, In: Theory of Automata*, chapter II, pages 71–93. Pergamon Press, 1969.