Egyetemi doktori (PhD) értekezés tézisei

NEM HAGYOMÁNYOS MODELLEK A FORMÁLIS NYELVEK ÉS AUTOMATÁK ELMÉLETÉBEN

Hegedüs László

Témavezető: Dr. Nagy Benedek

DEBRECENI EGYETEM

Informatikai Tudományok Doktori Iskola

Debrecen, 2016

1. Bevezetés

Az utóbbi néhány évtizedben a számítástudomány területén egyre nagyobb figyelmet szentelnek a nem hagyományos modellek vizsgálatának. A fogalom igen tág, általában minden olyan modellt ide sorolhatunk, amely a klasszikus szakirodalomban [Hopcroft et al., 2006] tárgyaltaktól valamilyen formában eltér. A felmerülő ötletek egyik fő forrása a biológia. Ezen terület központi eleme a bioinspirált számítástudomány (bioinspired computing), mely az élőlényekben végbemenő biológiai és kémiai folyamatokon alapuló műveletek elméleti és gyakorlati alkalmazási lehetőségeit vizsgálja. Kutatásaim során két bioinspirált modellel foglalkoztam.

A körszavak biológiai motivációjaként tekinthetünk a gyűrű alakú DNS láncokra (például a mitokondriális DNS [Kiss, 2007]). Ugyanakkor, az általam alkalmazott megközelítés matematikai eszközöket használ és a biológiai motivációtól függetlenül foglalkozik a körszavak bizonyos tulajdonságaival. A körszó, mint modell nem újdonság a szakirodalomban [Currie and Fitzpatrick, 2002, Diekert et al., 2006, Shur, 2010], viszont az általam vizsgált periódus fogalom új, még nem ismert eredmények felfedezéséhez vezetett.

Kutatásaim másik tárgyát képezik az egyállapotú $5' \to 3'$

Watson–Crick számlálóautomaták, melyek a DNS számítástudomány területén erősen kutatott Watson–Crick automaták speciális változatai. Az egyállapotú $5' \rightarrow 3'$ Watson–Crick számlálóautomata egy új modell, amelynek különböző paraméterei jelentősen befolyásolják kifejezőerejét. Eredményeim leginkább az így kialakuló hierarchia feltárásán alapulnak.

2. Célkitűzések és eredmények

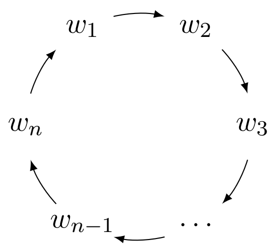
2.1. Körszavak kombinatorikájával kapcsolatos célkitűzések

A szavak kombinatorikus tulajdonságainak elemzése egészen az 1900-as évek elejére nyúlik vissza, mikor is *Axel Thue* publikálta a négyzetmentes szavakkal kapcsolatos műveit [Thue, 1906, 1912]. Az első áttekintő, összefoglaló művet *M. Lothaire* írói álnév alatt publikáló matematikusok csoportja írta [Lothaire, 1997]. Legtöbbjük *Marcel P. Schützenberger* tanítványa volt, akinek a nevéhez több fontos tétel is kapcsolódik. Valamennyi itt említett mű leginkább a hagyományos (lineáris) szavak vizsgálatával foglalkozik, a körszavakkal csak elvétve találkozhatunk bennük, pedig igen érdekes megoldandó kérdések merülhetnek fel velük kapcsolatban. Célkitűzéseim és eredményeim közléséhez néhány alapfogalom áttekintése szükséges. Feltételezem a formális nyelvek és automaták egy bevezető kurzuson elhangzó alapfogalmainak ismeretét [Dömösi et al., 2011, Hopcroft et al., 2006].

Adott Σ ábécé feletti $\boldsymbol{u}$ és $\boldsymbol{v}$ szavak pontosan akkor ***egymás konjugáltjai*** ($\boldsymbol{u} \sim \boldsymbol{v}$), ha léteznek olyan $\boldsymbol{x}, \boldsymbol{y} \in \Sigma^*$ szavak, hogy $\boldsymbol{u} = \boldsymbol{xy}$ és $\boldsymbol{v} = \boldsymbol{yx}$. Legyen $\boldsymbol{w} \in \Sigma^*$, ekkor a $\boldsymbol{w}$−ből képzett körszó alatt a

$$\boldsymbol{w}_\circ = \{\boldsymbol{v} \in \Sigma^* \mid \boldsymbol{v} \sim \boldsymbol{w}\}$$

halmazt értjük.



1. ábra. Körszó

Körszavakkal kapcsolatos a ciklikus elforgatás művelete is, melyet $\sigma(.)$−val jelölünk és, ha $\boldsymbol{w} = w_1 w_2 \cdots w_n$, akkor $\sigma(\boldsymbol{w}) = w_2 \cdots w_n w_1$. Továbbá, $\sigma^0(\boldsymbol{w}) = \boldsymbol{w}$ és, ha $k \in \mathbb{N}$, akkor $\sigma^k(\boldsymbol{w}) = \sigma(\sigma^{k-1}(\boldsymbol{w}))$.

Legyen $\boldsymbol{f}_0 = b$, $\boldsymbol{f}_1 = a$ és tetszőleges $n \geq 2$ esetén $\boldsymbol{f}_n = \boldsymbol{f}_{n-1}\boldsymbol{f}_{n-2}$. A $\boldsymbol{f}_i$ szót ($i \in \mathbb{N}_0$) az $i-$edik ***Fibonacci szónak*** [Smyth, 2003], a belőle képzett körszót pedig $i-$edik ***Fibonacci körszónak*** nevezzük. Ezen szavak hossza pontosan a Fibonacci számok sorozatát követi. Az $\{\boldsymbol{f}_i\}_{i=0}^{\infty}$ rekurzív sorozat határértékét szokás nevezni ***végtelen Fibonacci szónak***.

A körszavak modelljével kapcsolatban a következő feladatok megoldását, kérdések megválaszolását tűztem ki célul.

1. Azt mondjuk, hogy a $\boldsymbol{w}$ szónak a $p > 0$ egész szám ***periódusa***, ha valamennyi $i = 1, \ldots, |\boldsymbol{w}| - p$ esetén teljesül $w_i = w_{i+p}$. Hogyan általánosítható körszavakra a hagyományos szavakon így definiált periódus fogalma?

2. Létezik-e az így általánosított periódus fogalom esetén adott periódus meglétére vonatkozó szükséges és elégséges feltétel?

3. *Nathan J. Fine* és *Herbert S. Wilf* periodicitási lemmája a hagyományos szavak kombinatorikájának egyik legfontosabb eredményei közé sorolható [Lothaire, 1997]. Adott szó két tetszőleges periódusa közötti kapcsolatot írja le, mégpedig azt mondja ki, hogy, ha p és q is a $\boldsymbol{w}$ szó periódusa, valamint $p+q-\text{lnko}(p,q) \leq |\boldsymbol{w}|$, akkor $\text{lnko}(p,q)$ is a $\boldsymbol{w}$ periódusa. Lehetséges-e a Fine és Wilf periodicitási lemma körszavak periódusaira való általánosítása?

4. Hagyományos szavak zárószeleteit szokás hierarchikus adatszerkezetekkel ábrázolni [Crochemore and Rytter, 2002]. Hogyan használhatunk hasonló adatszerkezeteket körszavak ábrázolására?

2.2. Körszavak kombinatorikájával kapcsolatos eredmények

Itt ismertetem, hogy az előző fejezetben megadott célkitűzésekre milyen válaszok születtek. A hagyományos szavakon definiált periódus fogalmát felhasználva definiáltam a ***gyenge periódust***, amely a körszó, mint halmaz legalább egy elemének periódusa. Azaz p a $\boldsymbol{w}_\circ$ körszó gyenge periódusa pontosan akkor, ha létezik olyan $\boldsymbol{v} \in \boldsymbol{w}_\circ$, hogy p a $\boldsymbol{v}$ periódusa [6].

1. Tézis (Gyenge periódus létezésének szükséges és elégséges feltétele [1])**.** *Legyen $\boldsymbol{w} \in \Sigma^*$ egy tetszőleges szó. Ha p a $\boldsymbol{w}_\circ$ egy gyenge periódusa, akkor létezik olyan $|\boldsymbol{w}|-p$ hosszú $\boldsymbol{x} \in \Sigma^*$ szó és $\ell \in \mathbb{N}_0$, hogy $\sigma^\ell(\boldsymbol{x})\boldsymbol{x}$ a $\boldsymbol{ww}$ egy faktora.*

A következők is teljesülnek:

I. Ha $\boldsymbol{ww}$ tartalmaz egy olyan $\boldsymbol{xx}$ ($\boldsymbol{x} \in \Sigma^+$) faktort, melyre $|\boldsymbol{x}| < |\boldsymbol{w}|$, akkor $p = |\boldsymbol{w}| - |\boldsymbol{x}|$ a $\boldsymbol{w}_\circ$ egy gyenge periódusa.

II. Ha létezik olyan $p > 0$ egész, hogy a következő feltételek fennállnak:

- *$|\boldsymbol{w}| = k \cdot p + r$ valamely $k > 0$, $0 \leq r < p$ esetén, valamint*
- *létezik olyan p hosszú $\boldsymbol{v} \in \Sigma^*$ szó és annak egy $\boldsymbol{v'}$ kezdőszelete, hogy valamely $0 < s < k$ esetén a $\boldsymbol{ww}$ szó egy faktora*

$$\sigma^{s \cdot p}(\boldsymbol{v}^{k-1}\boldsymbol{v'})\sigma^{(s-1)\cdot p}(\boldsymbol{v}^{k-1}\boldsymbol{v'}),$$

akkor p a $\boldsymbol{w}_\circ$ körszó gyenge periódusa. ■

Első tézisem tehát kimondja, hogy egy adott körszó milyen feltételek teljesülése mellett rendelkezik adott gyenge periódussal.

2. Tézis (Relatív prím gyenge periódusok [1])**.** *Legyenek $p, q \in \mathbb{N}$ olyanok, hogy $2 \leq q < p$. Továbbá, legyen $\boldsymbol{w} \in \Sigma^+$ egy olyan szó, hogy $\boldsymbol{w}_\circ$ gyenge periódusa p és q. Jelölje m a $(p \bmod q)$ értéket. Ha $\mathrm{lnko}(p, q) = 1$ és $\boldsymbol{w}$ nem unáris szó, akkor $|\boldsymbol{w}|$–ra teljesül a következő egyenlőtlenség:*

$$|\boldsymbol{w}| \leq \begin{cases} 2p - m\,, & \text{ha } m = 1, \\ 2p + q - m\,, & \text{ha } m > 1. \end{cases}$$

■

Ez az eredmény Fine és Wilf periodicitási lemmájának relatív prím gyenge periódusokra vonatkozó esetének általánosítása. A lemma teljes általánosítása még nyitott probléma, de ez az eredmény jelentősen hozzájárul a megoldásához. További kapcsolódó eredményként tekinthetjük az alábbi tézist, amely adott hosszú, nem unáris körszó különböző gyenge periódusainak maximális számára ad pontos felső korlátot.

3. Tézis (Különböző gyenge periódusok maximális száma [1])**.** *Legyen $n \geq 3$. Egy $\{a, b\}$ ábécé feletti n hosszú nem unáris szóból képzett körszó különböző gyenge periódusainak száma legfeljebb*

$$\begin{array}{ll} \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n}{4} \right\rfloor + 1, & \textit{ha } n \textit{ páratlan,} \\ \frac{n}{2} + 1, & \textit{ha } n \textit{ páros és } (n \; mod \; 3) = 2, \\ \frac{n}{2}, & \textit{egyébként.} \end{array}$$

■

Körszavak ábrázolásával kapcsolatban a fa adatszerkezetben rejlő lehetőségeket vizsgáltam [1, 4]. A zárószeletfából (suffix trie) kiindulva bevezettem a körszóhoz tartozó fa fogalmát, amelynek a gyökérből levélelemekbe vezető útjain a körszó egyes elemei állnak elő. Ezt formálisan a következő definícióval adtam meg.

1. Definíció. *Azt mondjuk, hogy a $\tau_{\boldsymbol{w}_\circ}$ a $\boldsymbol{w}_\circ$ körszó fája, ha $v = v_1 \cdots v_n \in \boldsymbol{w}_\circ$ pontosan akkor, ha léteznek $\tau_{\boldsymbol{w}_\circ}$–ben olyan $t_0, \ldots, t_n$ csomópontok, hogy*

- *t_0 a $\tau_{\boldsymbol{w}_\circ}$ fa gyökere, t_n levélelem, és*
- *minden $i = 1, \ldots, n$ esetén létezik a t_{i-1} csomópontból t_i csomópontba vezető él, melynek címkéje v_i.*

Fákkal való ábrázolás segítségével a Fibonacci szavakból képzett körszavak esetén ismétlődő mintát figyeltem meg, mely a következő tézis alapjául szolgált.

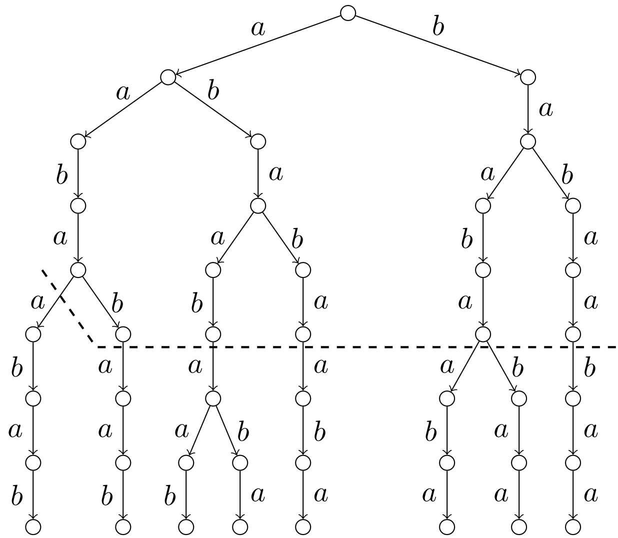
4. Tézis (Fibonacci körszavakhoz tartozó fák hierarchiája [1, 4])**.** *Bármely $2 < i < j$ esetén, $\tau_{(\boldsymbol{f}_i)_\circ}$ gyökérből kiinduló részfája a $\tau_{(\boldsymbol{f}_j)_\circ}$ fának.* ■

Azaz a rövidebb Fibonacci körszavak fái gyökérből kiinduló részfaként szerepelnek a hosszabb Fibonacci körszavak fáiban. Példaként a 2. ábrát emelem ki. Ennek következményeként a végtelen Fibonacci szóról a következőt állapítottam meg.

5. Tézis (A végtelen Fibonacci szó bizonyos faktorairól [1, 4])**.** *A végtelen Fibonacci szónak valamennyi $(\boldsymbol{f}_i)_\circ$−beli szó faktora bármely $i \in \mathbb{N}_0$ esetén.* ■

Ez alátámasztja, hogy a modell vizsgálata a hagyományos modellek szempontjából sem elhanyagolható.

A fent említett fontosabb eredmények mellett a bizonyításukhoz szükséges segédtételeket és azok bizonyításait is tartalmazza a dolgozat.



2. ábra. Látható, hogy $\tau_{(\boldsymbol{f}_5)_\circ}$ a $\tau_{(\boldsymbol{f}_6)_\circ}$ fa gyökérből kiinduló részfája. A $\tau_{(\boldsymbol{f}_5)_\circ}$ fa határát szaggatott vonal jelöli.

2.3. Egyállapotú $5' \to 3'$ Watson-Crick számlálóautomatákkal kapcsolatos célkitűzések

Számlálóautomatákkal, mint kiszámíthatósági modellekkel is igen régóta lehet találkozni a szakirodalomban. A témában az egyik legjelentősebb mű *Marvin L. Minsky* nevéhez fűződik, aki megmutatta, hogy a két számlálóval rendelkező számlálóautomatákkal valamennyi algoritmizálható probléma megoldását szimulálni lehet [Minsky, 1967]. *Ömer Eğecioğlu* és *Oscar H. Ibarra* munkái [Eğecioğlu and Ibarra, 2009a,b] ezen automaták egyállapotú (állapot nélküli, stateless) változataival foglalkoznak. Témavezetőm, *Nagy Benedek* végzett kutatásokat $5' \rightarrow 3'$ Watson–Crick automatákkal kapcsolatban, melyek a bioinspirált számítási modellek közé tartoznak. Munkáikra építve, *Ömer Eğecioğlu*val együtt dolgozva kezdtünk foglalkozni egyállapotú $5' \rightarrow 3'$ Watson–Crick számlálóautomatákkal, melyekkel kapcsolatban a következő kérdésekre, megoldandó problémákra kerestem a választ.

1. Hogyan írható le formálisan az egyállapotú $5' \rightarrow 3'$ Watson–Crick számlálóautomata modellje?
2. Ahogy az érzékelő $5' \rightarrow 3'$ Watson–Crick automatáknál [Nagy, 2008] is felmerült, elkerülendő az olvasófejek egymáson való "átlépése".
3. Hol helyezkedik el az új modell a Chomsky hierarchiában?
4. Az új modell különböző paraméterei alapján mit mondhatunk az elfogadott nyelvosztályok közötti hierarchiáról?

2.4. Egyállapotú $5' \to 3'$ Watson-Crick számlálóautomatákkal kapcsolatos eredmények

Itt ismertetem, hogy az előző fejezetben megadott célkitűzésekre milyen válaszok születtek. Az egyállapotú számlálóautomaták és több olvasófejjel rendelkező automaták definícióinak egységesítése után megadtam az egyállapotú $5' \to 3'$ Watson–Crick számlálóautomata definícióját a következőképpen.

2. Definíció ([8])**.** ***Egyállapotú*** $5' \to 3'$ ***Watson–Crick számlálóautomatának*** *nevezzük az* $\mathcal{A} = (\Sigma, ¢, \$, m, \delta)$ *ötöst, ahol*

- Σ *egy nem üres, véges* ***ábécé****,*
- $¢, \$ \notin \Sigma$ *rendre* ***kezdő*** *és* ***záró jelek****,*
- $m \in \mathbb{N}_0$ *az automata számlálóinak darabszáma,*
- $\delta = \delta_1 \cup \delta_2$ *az* ***átmenetfüggvény****, ahol*

$$\delta_1 : (\Sigma \cup \{¢, \$\})^2 \times \{0,1\}^m \to \mathcal{P}(\{0,1\}^2 \times \{0,+,-\}^m)$$

$$\delta_2 : \{\lambda\}^2 \times \{0,1\}^m \to \mathcal{P}(\{0\}^2 \times \{0,+,-\}^m),$$

továbbá megköveteljük, hogy tetszőleges $a, b \in (\Sigma \cup \{¢, \$\})$*, valamint* $s = (s_1, \ldots, s_m) \in \{0,1\}^m$ *esetén, ha* $\delta(a, b, s)$ *értelmezett (nem az üres halmaz), akkor nem létezik olyan* $i \in \{1, \ldots, m\}$*, hogy* $s_i = 0$ *és* $t_i = -$*, ahol* $(t_1, \ldots,$

$t_m) \in \delta(a, b, s)$, azaz nulla értékű számláló csökkentése nem megengedett.

A definícióból kiindulva a számítási folyamatot leíró konfigurációkat és a köztük való átmenetet úgy definiáltam, hogy az olvasófejek "találkozását" is figyelembe vesszük, azok nem léphetnek át egymáson.

A $C = (\boldsymbol{w}, s) \in \{¢, \lambda\}\Sigma^*\{\$, \lambda\} \times \mathbb{N}_0^m$ párt az $\mathcal{A}$ automata egy konfigurációjának nevezzük. Az egy olvasófejjel rendelkező automatákhoz hasonlóan a

$$C_2 = (\boldsymbol{w'}, s') = (\boldsymbol{w'}, s'_1, \ldots, s'_m)$$

konfiguráció közvetlenül elérhető a

$$C_1 = (\boldsymbol{w}, s) = (\boldsymbol{w}, s_1, \ldots, s_m)$$

konfigurációból (jelöléssel: $C_1 \vdash_{\mathcal{A}} C_2$), ha az alábbiak közül az egyik teljesül:

- $\boldsymbol{w} = a\boldsymbol{w'}b$, $|\boldsymbol{w}| \geq 2$ és $(1, 1, t_1, \ldots, t_m) \in \delta(a, b, \mathrm{sign}(s))$,
- $\boldsymbol{w} = a\boldsymbol{w'} = \boldsymbol{u}b$ valamely $\boldsymbol{u} \in \Sigma^*$ esetén, $|\boldsymbol{w}| \geq 1$ és $(1, 0, t_1, \ldots, t_m) \in \delta(a, b, \mathrm{sign}(s))$,
- $\boldsymbol{w} = \boldsymbol{w'}b = a\boldsymbol{u}$ valamely $\boldsymbol{u} \in \Sigma^*$ esetén, $|\boldsymbol{w}| \geq 1$ és $(0, 1, t_1, \ldots, t_m) \in \delta(a, b, \mathrm{sign}(s))$,
- $\boldsymbol{w} = a\boldsymbol{u} = \boldsymbol{v}b = \boldsymbol{w'}$ valamely $\boldsymbol{u}, \boldsymbol{v} \in \Sigma^*$ szavak esetén, $|\boldsymbol{w}| \geq 1$ és $(0, 0, t_1, \ldots, t_m) \in \delta(a, b, \mathrm{sign}(s))$,

- $\boldsymbol{w} = \lambda = \boldsymbol{w'}$ és $(0, 0, t_1, \ldots, t_m) \in \delta(\lambda, \lambda, \mathrm{sign}(s))$,

továbbá

$$s'_i = \begin{cases} s_i + 1\,, & \text{ha } \ t_i = +, \\ s_i - 1\,, & \text{ha } \ s_i > 0 \text{ és } t_i = -, \\ s_i\,, & \text{egyébként (azaz, ha } \ t_i = 0), \end{cases}$$

minden $i = 1, \ldots, m$–re.

Az $\mathcal{A}$ automata által elfogadott nyelv

$$\mathcal{L}(\mathcal{A}) = \{\boldsymbol{w} \in \Sigma^* \mid (¢\boldsymbol{w}\$, 0^m) \vdash^* (\lambda, 0^m)\}.$$

3. Definíció. *Egy egyállapotú $5' \to 3'$ Watson–Crick számlálóautomata* ***valós-idejű****, ha minden a, $b \in (\Sigma \cup \{¢, \$\})$, $s \in \{0,1\}^m$ esetén*

$$\delta(a, b, \mathrm{sign}(s)) \subseteq \left\{(d_1, d_2, t) \in \{0,1\}^2 \times \{0, +, -\}^m \mid d_1 + d_2 \geq 1\right\}.$$

Azaz, ha minden értelmezett átmenet esetén legalább az egyik olvasófej elmozdul. Ha ez a korlátozás nincs előírva, akkor az automatát ***nem valós-idejűnek*** *nevezzük.*

Egy egyállapotú $5' \to 3'$ Watson–Crick számlálóautomata ***determinisztikus***, ha valamennyi a, $b \in (\Sigma \cup \{¢, \$\})$ és $s \in \{0,1\}^m$ esetén $|\delta(a, b, \mathrm{sign}(s))| \leq 1$. Azaz egy egyállapotú $5' \to 3'$ Watson–Crick számlálóautomata pontosan akkor determinisztikus, ha minden konfiguráció esetén, amennyiben létezik lehetséges lépés (azaz az átmenetfüggvény nem az üres halmazt adja eredményül), akkor az egyértelműen meghatározott.

Legyen r egy számláló és vizsgáljuk a $C_1 \vdash \ldots \vdash C_\ell$ konfigurációk sorozatát, amelyet ***számításnak*** is nevezünk. Minden számítás során figyelemmel lehet követni, hogy az egyes számlálók hányszor kerülnek nemcsökkenő (növelések és változatlanul hagyások sorozata) állapotból csökkenőbe (legalább egy csökkentés művelet után). Ha ez adott r számlálóval tetszőleges számítás során legfeljebb $k-$szor történik, akkor azt mondjuk, hogy r $k-$szor fordul, vagy r egy $k-$***forduló számláló*** ($k-$reversal counter). Ha nem létezik ilyen k egész, akkor r fordulásainak száma korlátlan és ***korlátlanul fordulónak*** nevezzük. Megjegyezzük, hogy az elfogadó és nem elfogadó számításokat is figyelembe vesszük r fordulásainak meghatározásánál. Sokszor egyes számlálók csak arra szolgálnak, hogy számon tartsák más számlálók fordulásait. Így oldható meg az, hogy, ha például az r számláló fordulásainak száma elérte a k határértéket, akkor az automata elutasító konfigurációba kerüljön.

A számlálókhoz hasonlóan, ha létezik egy olyan k egész, hogy adott $\mathcal{A}$ automatának minden számlálója (legfeljebb) $k-$forduló, akkor az $\mathcal{A}$ automatát is $k-$***forduló automatá-***

nak nevezzük. Ellenkező esetben $\mathcal{A}$ ***korlátlanul forduló automata***.

Ahogy az alábbi tézis mutatja, beláttam, hogy létezik olyan reguláris nyelv, melyet véges sok számlálóval nem fogad el egyetlen determinisztikus valós-idejű egyállapotú $5' \rightarrow 3'$ Watson–Crick számlálóautomata sem.

6. Tézis (Determinisztikus számlálóautomaták korlátai [2]). *Az $L_{ab} = \{(ab)^n \mid n \in \mathbb{N}\}$ nyelvet egyetlen determinisztikus valós-idejű egyállapotú $k-$forduló $m-$számlálós automata sem fogadja el ($k, m \geq 1$).* ■

Megadtam olyan négy számlálóval rendelkező automata definícióját, amely elfogad egy nem környezetfüggetlen, nem szemilineáris nyelvet.

7. Tézis (Egy szemilineáris, nem környezetfüggő nyelvet elfogadó automata [2]). *Létezik olyan determinisztikus nem valós-idejű egyállapotú $5' \rightarrow 3'$ Watson–Crick számlálóautomata, amely az $L = \{a^{n^2} \mid n \geq 0\}$ nyelvet fogadja el.* ■

Melynek következménye, hogy a determinisztikus nem valós-idejű egyállapotú $5' \rightarrow 3'$ Watson–Crick számlálóautomaták által elfogadott nyelvek osztálya tartalmaz nem szemilineáris, így nem környezetfüggetlen nyelvet.

8. Tézis (Nemdeterminisztikus automaták kifejezőereje [7]). *Létezik olyan nyelv, melyet elfogad egy nemdeterminisztikus valós-idejű egyállapotú* $1-$*forduló* $1-$*számlálós automata, de egyetlen*
$k, m \in \mathbb{N}$ *esetén sem fogadja el egy determinisztikus valós-idejű egyállapotú* $k-$*forduló* $m-$*számlálós automata.* ■

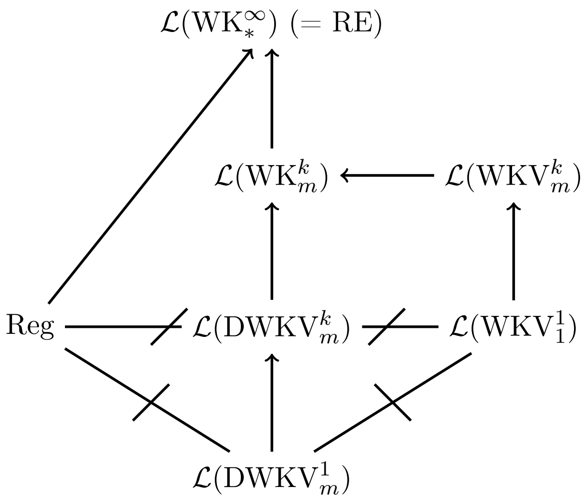
Azaz az egyállapotú $5' \to 3'$ Watson–Crick számlálóautomaták nemdeterminisztikus változatainak kifejezőereje azonos paraméterek mellett nagyobb, mint a determinisztikus változatoké.

Eredményeimet a 3. ábrán látható tartalmazási hierarchiával lehet ábrázolni. Értelmezéséhez az alábbi jelölések bevezetése szükséges. Legyenek $k, m \in \mathbb{N}$ és $k < \frac{2^{m-1}}{m}$, ugyanis az automata implicite oldja meg a számlálók fordulásainak korlátozását. Jelölje WK_m^k a $k-$forduló $m-$számlálós (nemdeterminisztikus) nem valós-idejű automaták osztályát, DWK_m^k pedig az azonos paraméterekkel rendelkező determinisztikus automatákból álló osztályt. Hasonlóan, rendre WKV_m^k és DWKV_m^k jelöli a valós-idejű, valamint a determinisztikus valós-idejű részhalmazokat. Továbbá bevezetjük a WK_m^∞ jelölést, amely a korlátozás nélkül forduló m-számlálós automatákat jelöli és legyen

$$\mathrm{WK}_*^\infty = \bigcup_{m\in\mathbb{N}} \mathrm{WK}_m^\infty.$$

Ha K egy automataosztály, akkor $\mathcal{L}(K)-$val jelöljük az osztályba tartozó automaták által elfogadott nyelveket, azaz

$$\mathcal{L}(K) = \{\mathcal{L}(\mathcal{A}) \mid \mathcal{A} \in K\}.$$



3. ábra. Az ábrán Reg jelöli a reguláris, RE pedig a rekurzív felsorolható nyelvek osztályát. A nyilak valódi tartalmazást jelentenek, az áthúzott vonalak pedig azt jelzik, hogy a két nyelvosztály halmazelméleti tartalmazást tekintve nem összehasonlítható, azaz mindkét osztályban vannak olyan nyelvek, amelyek a másikban nem szerepelnek. Ha két osztály között az ábrán nincs jelölve semmilyen reláció, akkor annak teljes felderítése még megoldatlan feladat. Továbbá a tartalmazási reláció tranzitivitása miatt a redundáns nyilakat elhagytam.

Theses of the PhD dissertation

UNCONVENTIONAL MODELS IN THE THEORY OF FORMAL LANGUAGES AND AUTOMATA

László Hegedüs

Supervisor: Dr. Benedek Nagy

UNIVERSITY OF DEBRECEN
Doctoral School of Informatics
2016, Debrecen

1 Introduction

During the last decades we can observe a growing interest in unconventional models of computation. The field is vast, since it consists of all models that are different in one way or another from the ones described in classical technical literature [Hopcroft et al., 2006]. Biology proved to be a major source of ideas. The central element of the field is bioinspired computing which deals with the practical and theoretical applications of biological and chemical processes observed inside living things. My research is based on two models.

In nature we can find rings formed from DNA strands (e.g., mitochondrial DNA) that can be considered as a motivation for analyzing the properties of circular words. Nevertheless, my approach to circular words is purely mathematical and independent of the biological motivation. The notion of circular word is not new, other researchers investigated certain properties of them [Currie and Fitzpatrick, 2002, Diekert et al., 2006, Shur, 2010], but I introduced a new notion of periodicity which lead to the discovery of brand new results.

The other topic of my research is the stateless multicounter $5' \to 3'$ Watson–Crick automaton which is a special variation of the well known Watson–Crick automaton of DNA computing.

The model is new and it has several parameters that affect the expressing power significantly. Therefore my research targeted the hierarchy of accepted languages based on these parameters.

2 Primary goals and results

2.1 Primary goals regarding periodic properties of circular words

The analysis of periodic properties of words goes back to the beginning of the 1900s when *Axel Thue* published his articles about square-free words [Thue, 1906, 1912]. The first compilation of results in this area was written by a group of people under the pen name *M. Lothaire* [Lothaire, 1997]. Most of them were students of *Marcel P. Schützenberger* who is well known for his contributions to this field. All of the work mentioned above deals with ordinary (linear) words and rarely mention circular words, even though the field has great potential and many interesting open questions. To introduce my goals and results regarding the topic of circular words a couple of notions have to be defined. I assume that the reader is familiar with the basic notions and notation of the theory of formal languages and automata [Hopcroft et al., 2006]. An introductory course should suffice. Two words $\boldsymbol{u}$ and $\boldsymbol{v}$ over an alphabet Σ are ***conjugates*** if and only if there exist $\boldsymbol{x}, \boldsymbol{y} \in \Sigma^*$ such that $\boldsymbol{u} = \boldsymbol{xy}$ and $\boldsymbol{v} = \boldsymbol{yx}$. Let $\boldsymbol{w} \in \Sigma^*$. Then the circular word obtained from $\boldsymbol{w}$ is the set of words

$$\boldsymbol{w}_\circ = \{\boldsymbol{v} \in \Sigma^* \mid \boldsymbol{v} \sim \boldsymbol{w}\}.$$

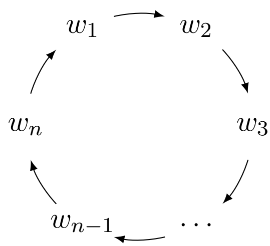


Figure 1. Circular word

In connection to circular words there exists the operator $\sigma(.)$ of cyclic shift. If $\boldsymbol{w} = w_1 w_2 \cdots w_n$, then

$$\sigma(\boldsymbol{w}) = w_2 \cdots w_n w_1.$$

Furthermore, $\sigma^0(\boldsymbol{w}) = \boldsymbol{w}$ and, if $k \in \mathbb{N}$, then

$$\sigma^k(\boldsymbol{w}) = \sigma(\sigma^{k-1}(\boldsymbol{w})).$$

Let $\boldsymbol{f}_0 = b$, $\boldsymbol{f}_1 = a$ and for all $n \geq 2$ $\boldsymbol{f}_n = \boldsymbol{f}_{n-1}\boldsymbol{f}_{n-2}$. The word $\boldsymbol{f}_i$ $(i \in \mathbb{N}_0)$ is called the ith ***Fibonacci word*** [Smyth, 2003], and the circular word obtained from it is the ith ***circular Fibonacci word***. The length of these words follow exactly the series of Fibonacci numbers. The limit of the series $\{\boldsymbol{f}_i\}_{i=0}^{\infty}$ is usually referred to as the ***infinite Fibonacci word***.

Regarding the model of circular words the goals of my work was to investigate and give answers to the following questions.

1. We say that an integer $p > 0$ is a ***period*** of the word $\boldsymbol{w}$ if and only if $w_i = w_{i+p}$ for all $i = 1, \ldots, |\boldsymbol{w}| - p$. How can the notion of period be generalized to circular words?
2. Is it possible to give a necessary and sufficient condition on the existence of a given generalized period?
3. The Periodicity Lemma of *Nathan J. Fine* and *Herbert S. Wilf* is one of the most important results in combinatorics on words [Lothaire, 1997]. It describes the relation of two arbitrary periods of a given word and states that, if both p and q are periods of the word $\boldsymbol{w}$ and $p + q - \mathrm{lnko}(p, q) \leq |\boldsymbol{w}|$, then $\mathrm{lnko}(p, q)$ is also a period of $\boldsymbol{w}$. Is it possible to generalize this result to periods of circular words?
4. Suffixes of ordinary words are often represented by suffix tries [Crochemore and Rytter, 2002]. How can we make use of similar hierarchical data structures in the repre-

sentation of circular words?

2.2 Results regarding periodic properties of circular words

Here I present the results of my work regarding the problems and questions mentioned above. Using the notion of period for ordinary words I defined the ***weak period*** which is a period of at least one element of a given circular word. That is, the integer $p > 0$ is a weak period of the circular word $\boldsymbol{w}_\circ$ if and only if p is a period of at least one $\boldsymbol{v} \in \boldsymbol{w}_\circ$ [6].

I found and proved a necessary and sufficient condition for a given integer p to be a weak period of a given circular word $\boldsymbol{w}_\circ$ [1]. This result is given by the following Thesis.

Thesis 1 (A necessary and sufficient condition for the existence of a certain weak period [1])**.** *Let $\boldsymbol{w} \in \Sigma^*$ be an arbitrary word. If p is a weak period of $\boldsymbol{w}_\circ$, then there exists a word $\boldsymbol{x} \in \Sigma^*$ of length $|\boldsymbol{w}| - p$ and an integer $\ell \in \mathbb{N}_0$, such that $\sigma^\ell(\boldsymbol{x})\boldsymbol{x}$ is a factor of $\boldsymbol{ww}$.*

Conversely, the following are true:

I. If $\boldsymbol{ww}$ has a factor $\boldsymbol{xx}$ ($\boldsymbol{x} \in \Sigma^+$), such that $|\boldsymbol{x}| < |\boldsymbol{w}|$, then $p = |\boldsymbol{w}| - |\boldsymbol{x}|$ is a weak period of $\boldsymbol{w}_\circ$.

II. If there exists an integer $p > 0$, such that the following *are satisfied*

- $|\boldsymbol{w}| = k \cdot p + r$ *for some* $k > 0$, $0 \leq r < p$, *furthermore,*
- *there exists a word* $\boldsymbol{v} \in \Sigma^*$ *of length* p *with some prefix* $\boldsymbol{v'}$, *such that*

$$\sigma^{s \cdot p}(\boldsymbol{v}^{k-1}\boldsymbol{v'})\sigma^{(s-1) \cdot p}(\boldsymbol{v}^{k-1}\boldsymbol{v'}),$$

is a factor of $\boldsymbol{ww}$ *for some* $0 < s < k$, *then*

p *is a weak period of* $\boldsymbol{w}_\circ$. ■

Thesis 2 (Relatively prime weak periods [1])**.** *Let* $p, q \in \mathbb{N}$ *with* $2 \leq q < p$. *Moreover, let* $\boldsymbol{w} \in \Sigma^+$, *such that both* p *and* q *are weak periods of* $\boldsymbol{w}_\circ$. *Let* m *denote the value of* $(p \bmod q)$. *If lnko*$(p, q) = 1$ *and* $\boldsymbol{w}$ *is not a unary word, then*

$$|\boldsymbol{w}| \leq \begin{cases} 2p - m\,, & \text{if } m = 1, \\ 2p + q - m\,, & \text{if } m > 1. \end{cases}$$

■

The result above is a generalization of the case of the Periodicity Lemma of Fine and Wilf where p and q are relatively prime. The complete generalization of the Lemma remains an open problem, but I believe that my result covers a significant part of the solution. In connection to this issue I stated and proved a strict upper limit on the number of distinct weak periods of a non-unary circular word of given length.

Thesis 3 (Strict upper limit on the maximal number of distinct weak periods [1])**.** *Let $n \geq 3$. The maximal number of distinct weak periods of a non unary circular word of length n over the alphabet $\{a, b\}$ is at most*

$$\begin{array}{ll} \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n}{4} \right\rfloor + 1, & \text{if } n \text{ is odd,} \\ \frac{n}{2} + 1, & \text{if } n \text{ is even and } (n \bmod 3) = 2, \\ \frac{n}{2}, & \text{otherwise.} \end{array}$$

■

Regarding the representation of circular words I analyzed the potential in the tree data structure [1, 4]. Similarly to suffix tries I defined the notion tree of a circular word. A tree of a circular word is a hierarchical data structure such that there is a one-to-one correspondence between the routes from the root to the leaf nodes and the elements of the circular word. I gave the following definition.

Definition 1. *A tree $\tau_{\boldsymbol{w}_\circ}$ is the tree of the circular word $\boldsymbol{w}_\circ$ if and only if for all $v = v_1 \cdots v_n \in \boldsymbol{w}_\circ$ there exist nodes $t_0, \ldots,$ t_n in $\tau_{\boldsymbol{w}_\circ}$ such that*

- *t_0 is the root node of $\tau_{\boldsymbol{w}_\circ}$, t_n is a leaf node and*

- *for all $i = 1, \ldots, n$ there exists an edge with label v_i leading from t_{i-1} to t_i.*

This representation proved to be useful when I considered circular words obtained from Fibonacci words. I observed a recurring pattern in the trees of circular Fibonacci words and discovered the following result.

Thesis 4 (Hierarchy of the trees of circular Fibonacci words [1, 4])**.** *For all $2 < i < j$, $\tau_{(\boldsymbol{f}_i)_\circ}$ is a rooted subtree of $\tau_{(\boldsymbol{f}_j)_\circ}$.* ■

That is, the tree of a circular Fibonacci word contains all trees of shorter circular Fibonacci words as rooted subtrees. See Figure 2 for example. As a corollary I stated the following.

Thesis 5 (Certain factors of the infinite Fibonacci word [1, 4])**.** *For all $i \in \mathbb{N}_0$, all elements of $(\boldsymbol{f}_i)_\circ$ is a factor of the infinite Fibonacci word.* ■

This confirms that the analysis of this model should not be neglected even from the point of view of classical models.

Supplementing the results above, my dissertation contains the necessary lemmas with their proofs that build the foundation on which these results were based.

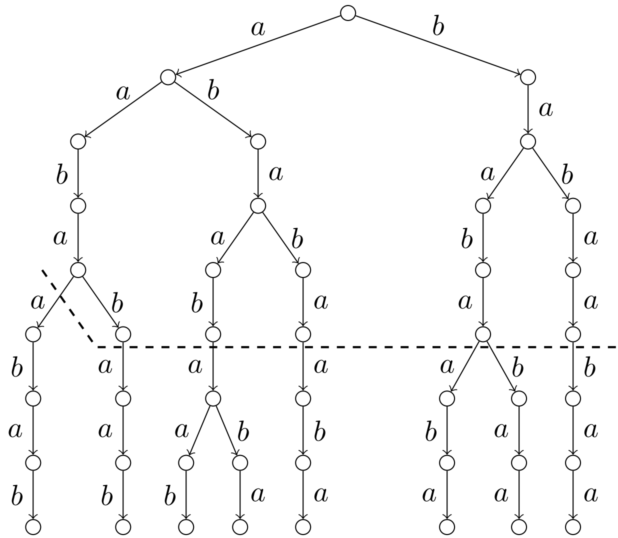


Figure 2. How $\tau_{(\boldsymbol{f}_5)_\circ}$ appears as a rooted subtree in $\tau_{(\boldsymbol{f}_6)_\circ}$. The border of $\tau_{(\boldsymbol{f}_5)_\circ}$ is denoted by a dashed line.

2.3 Goals regarding stateless multicounter $5' \rightarrow 3'$ Watson-Crick automata

The counter machine as a basis of computation can be found in classical literature. One of the most influential works in this area were written by *Marvin L. Minsky* who proved that the solution of any algorithmic problem can be simulated by counter machines that have only two counters [Minsky, 1967].

Some articles of *Ömer Eğecioğlu* and *Oscar H. Ibarra* deal with a special case of these automata, namely stateless multicounter machines [Eğecioğlu and Ibarra, 2009a,b]. My advisor, Benedek Nagy have done research on $5' \to 3'$ Watson–Crick automata which is a bioinspired computational model. Based on the previous research we started working together with *Ömer Eğecioğlu* on stateless multicounter $5' \to 3'$ Watson–Crick automata. During this work I was looking for answers to the following problems.

1. Give a formal definition of stateless multicounter $5' \to 3'$ Watson–Crick automata.
2. The reading heads of sensing $5' \to 3'$ Watson–Crick automata [Nagy, 2008] do not cross each-other during the computation.

   Define the behavior of stateless multicounter $5' \to 3'$ Watson–Crick automata to reflect this method of computation.
3. Where can we place the new model in the Chomsky-hierarchy?
4. What can we say about the classes of accepted languages based on the parameters of the model?

2.4 Results regarding stateless multicounter $5' \to 3'$ Watson-Crick automata

Here I present the results of my work regarding the problems and questions mentioned above.

After unifying the notation of stateless multicounter machines and multihead automata I gave the formal definition of stateless multicounter $5' \to 3'$ Watson–Crick automata.

Definition 2 ([8]). *A* ***stateless multicounter*** $5' \to 3'$ ***Watson–Crick automaton*** *is a quintuple* $\mathcal{A} = (\Sigma, ¢, \$, m, \delta)$, *where*

- Σ *is a non empty, finite* ***alphabet***,
- $¢, \$ \notin \Sigma$ *are respectively the* ***start*** *and* ***end*** *symbols,*
- $m \in \mathbb{N}_0$ *is the number of counters of the automaton,*
- $\delta = \delta_1 \cup \delta_2$ *is the* ***transition function***, *where*

$$\delta_1 : (\Sigma \cup \{¢, \$\})^2 \times \{0,1\}^m \to \mathcal{P}(\{0,1\}^2 \times \{0,+,-\}^m)$$

$$\delta_2 : \{\lambda\}^2 \times \{0,1\}^m \to \mathcal{P}(\{0\}^2 \times \{0,+,-\}^m),$$

furthermore, we require that for any $a, b \in (\Sigma \cup \{¢, \$\})$ *and* $s = (s_1, \ldots, s_m) \in \{0,1\}^m$, *if* $\delta(a, b, s)$ *is defined (not the empty set), then there does not exist* $i \in \{1, \ldots, m\}$, *such that* $s_i = 0$ *and* $t_i = -$, *where* $(t_1, \ldots,$

$t_m) \in \delta(a, b, s)$, that is, it is not possible to decrease a counter whose value is zero.

I defined the configurations and a relation over configurations that describe the computation of an automaton in a way that prohibits the reading heads to cross each-other.

The pair $C = (\boldsymbol{w}, s) \in \{¢, \lambda\}\Sigma^*\{\$, \lambda\} \times \mathbb{N}_0^m$ is a configuration of the automaton $\mathcal{A}$. The configuration

$$C_2 = (\boldsymbol{w'}, s') = (\boldsymbol{w'}, s'_1, \ldots, s'_m)$$

is directly accessible from configuration

$$C_1 = (\boldsymbol{w}, s) = (\boldsymbol{w}, s_1, \ldots, s_m)$$

(denoted by: $C_1 \vdash_{\mathcal{A}} C_2$), if and only if one of the followings is satisfied:

- $\boldsymbol{w} = a\boldsymbol{w'}b$, $|\boldsymbol{w}| \geq 2$ and $(1, 1, t_1, \ldots, t_m) \in \delta(a, b, \text{sign}(s))$,
- $\boldsymbol{w} = a\boldsymbol{w'} = \boldsymbol{u}b$ for some $\boldsymbol{u} \in \Sigma^*$, $|\boldsymbol{w}| \geq 1$ and $(1, 0, t_1, \ldots, t_m) \in \delta(a, b, \text{sign}(s))$,
- $\boldsymbol{w} = \boldsymbol{w'}b = a\boldsymbol{u}$ for some $\boldsymbol{u} \in \Sigma^*$, $|\boldsymbol{w}| \geq 1$ and $(0, 1, t_1, \ldots, t_m) \in \delta(a, b, \text{sign}(s))$,
- $\boldsymbol{w} = a\boldsymbol{u} = \boldsymbol{v}b = \boldsymbol{w'}$ for some $\boldsymbol{u}, \boldsymbol{v} \in \Sigma^*$, $|\boldsymbol{w}| \geq 1$ and $(0, 0, t_1, \ldots, t_m) \in \delta(a, b, \text{sign}(s))$,
- $\boldsymbol{w} = \lambda = \boldsymbol{w'}$ and $(0, 0, t_1, \ldots, t_m) \in \delta(\lambda, \lambda, \text{sign}(s))$,

furthermore,

$$s_i' = \begin{cases} s_i + 1\,, & \text{if } \; t_i = +, \\ s_i - 1\,, & \text{if } \; s_i > 0 \text{ and } t_i = -, \\ s_i\,, & \text{otherwise (that is, if } \; t_i = 0), \end{cases}$$

for all $i = 1, \ldots, m$.

The language accepted by $\mathcal{A}$ is defined as

$$\mathcal{L}(\mathcal{A}) = \{\boldsymbol{w} \in \Sigma^* \mid (¢\boldsymbol{w}\$, 0^m) \vdash^* (\lambda, 0^m)\}.$$

Definition 3. *A stateless multicounter* $5' \to 3'$ *Watson–Crick automaton is **realtime** if and only if for all* a, $b \in (\Sigma \cup \{¢, \$\})$, $s \in \{0,1\}^m$ *the following is true*

$$\delta(a, b, \mathrm{sign}(s)) \subseteq \left\{(d_1, d_2, t) \in \{0,1\}^2 \times \{0, +, -\}^m \mid d_1 + d_2 \geq 1\right\}.$$

*That is, at least one of the reading heads move in each transition. If this condition is not required to be satisfied, then we call the automaton **non-realtime**.*

A stateless multicounter $5' \to 3'$ Watson–Crick automaton is ***deterministic*** if and only if for all a, $b \in (\Sigma \cup \{¢, \$\})$ and $s \in \{0,1\}^m$, $|\delta(a, b, \mathrm{sign}(s))| \leq 1$.

Let r be a counter and $C_1 \vdash \ldots \vdash C_\ell$ a series of configurations which is also called a computation. The behavior of r can be observed during any computation regarding the changes between non-decreasing and decreasing phases, i.e., when the value of the counter is decreased after several steps of increase and unchange operations. If this happens to r at most k times in any computation, then we say that r is a $k-$reversal counter. If no such k exists, then r is a reversal unbounded counter.

If there exists an integer k, such that all counters of $\mathcal{A}$ are (at most) $k-$reversal, then $\mathcal{A}$ is a $k-$reversal automaton. Otherwise, $\mathcal{A}$ is a reversal unbounded automaton.

As the following thesis shows, I proved that there exists a regular language that cannot be accepted by any finite number of counters by a deterministic stateless multicounter $5' \to 3'$ Watson–Crick automaton.

Thesis 6 (Limits of deterministic stateless multicounter $5' \to 3'$ Watson–Crick automata [2])**.** *The language $L_{ab} = \{(ab)^n \mid n \in \mathbb{N}\}$ cannot be accepted by any deterministic realtime $k-$reversal $m-$counter $5' \to 3'$ Watson–Crick automaton for any $k, m \geq 1$.* ■

I also defined a deterministic stateless multicounter $5' \to 3'$ Watson–Crick automata with four counters that accepts a non-context-free semilinear language.

Thesis 7 (Accepting a non-context-free semilinear language [2])**.** *There exists a deterministic non-realtime stateless mul-*

ticounter $5' \to 3'$ Watson–Crick automaton that accepts the language $L = \{a^{n^2} \mid n \geq 0\}$. ■

Thus the class of languages accepted by deterministic non-realtime stateless multicounter $5' \to 3'$ Watson–Crick automata contains a semilinear non-context-free language.

Thesis 8 (The expressive power of nondeterministic automata [7])**.** *There exists a language that can be accepted by a nondeterministic realtime $1-$reversal $1-$counter automata, but cannot be accepted by any deterministic realtime $k-$reversal $m-$counter automaton for any $k, m \in \mathbb{N}$.*

That is, the expressive power of nondeterministic stateless multicounter $5' \to 3'$ Watson–Crick automata is greater than that of their deterministic counterparts.

My results are summarized on Figure 3. Let $k, m \in \mathbb{N}$ and $k < \frac{2^{m-1}}{m}$, since the automata implicitly solves reversal boundedness. Let WK_m^k denote the class of non-realtime (nondeterministic) $k-$reversal $m-$counter automata, while DWK_m^k is the class of deterministic automata with the same parameters. Similarly, WKV_m^k and DWKV_m^k denotes respectively the realtime and deterministic realtime classes. Furthermore, we introduce the notation WK_m^{∞} that is the class of reversal unbounded $m-$counter automata and let

$$\mathrm{WK}_*^\infty = \bigcup_{m \in \mathbb{N}} \mathrm{WK}_m^\infty.$$

If K is a class of automata, then $\mathcal{L}(K)$ denotes the set of languages accepted by automata that belong to K, that is,

$$\mathcal{L}(K) = \{\mathcal{L}(\mathcal{A}) \mid \mathcal{A} \in K\}.$$

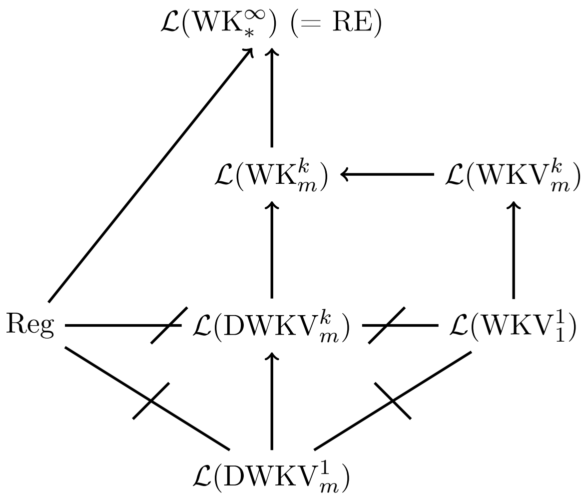


Figure 3. Reg and RE denotes the class of regular and recursive enumerable languages, respectively. Arrows denote strict inclusion, while strike-through lines denote that the two classes are incomparable (from a set theoretic point of view), that is, both classes contain at least one language that the other does not. If no relation is presented between two classes, then it is an open problem to describe their relation. The arrows (like the relations they denote) are transitive, thus the redundant ones are not drawn.

Irodalomjegyzék / References

Maxime Crochemore and Wojciech Rytter. ***Jewels of Stringology: Text Algorithms***. World Scientific, Hackensack, NJ, USA, 2002.

James D. Currie and Daniel S. Fitzpatrick. Circular words avoiding patterns. ***Proceedings of the 6th International Conference on Developments in Language Theory. LNCS***, 2450:319–325, 2002.

Volker Diekert, Tero Harju, and Dirk Nowotka. Factorizations of cyclic words. ***Workshop on Words and Automata at CSR***, 7, 2006.

Pál Dömösi, János Falucskai, Géza Horváth, Zoltán Mecsei, and Benedek Nagy. ***Formális nyelvek és automaták***. Debreceni Egyetem, jegyzet, 2011.

Ömer Eğecioğlu and Oscar H. Ibarra. On stateless multicounter machines. In ***Proceedings of 5th Conference on Computability in Europe 2009***, number 5635 in Lecture Notes in Computer Science, pages 178–187. Springer Berlin Heidelberg, 2009a.

Ömer Eğecioğlu and Oscar H. Ibarra. On stateless multicounter machines. In ***Proceedings of 15th International

Computing and Combinatorics Conference, number 5609 in International Computing and Combinatorics Conference, Lecture Notes in Computer Science, pages 408–417. Springer Berlin Heidelberg, 2009b.

John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. ***Introduction to Automata Theory, Languages, and Computation, 3rd ed***. Addison-Wesley, 2006.

János Kiss. ***Biológiai kislexikon***. Typotex Elektronikus Kiadó Kft, 2007.

M. Lothaire. ***Combinatorics on Words***. Cambridge Mathematical Library. Cambridge University Press (1983, Addison-Wesley), 1997.

Marvin L. Minsky. ***Computation: Finite and Infinite Machines***. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1967.

Benedek Nagy. On $5' \rightarrow 3'$ sensing Watson-Crick finite automata. In ***13th International Meeting on DNA Computing***, volume Lecture Notes in Computer Science, 4848 of ***International Meeting on DNA Computing***, pages 256–262. Springer Berlin Heidelberg, 2008.

Arseny M. Shur. On ternary square-free circular words. ***The Electronic Journal of Combinatorics***, 17, 2010.

William F. Smyth. ***Computing Patterns in Strings***. ACM Press. Pearson Addison-Wesley (UK), 2003.

Axel Thue. Über unendliche Zeichenreihen. ***Norske Vid. Skrifter I Mat.-Nat. Kl., Christiania***, 1:1–22, 1906.

Axel Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. ***Norske Vid. Skrifter I Mat.-Nat. Kl., Christiania***, 1:1–67, 1912.

Nyilvántartási szám:	DEENK/87/2016.PL
Tárgy:	PhD Publikációs Lista

Jelölt: Hegedüs László
Neptun kód: YU4N37
Doktori Iskola: Informatikai Tudományok Doktori Iskola
MTMT azonosító: 10037392

A PhD értekezés alapjául szolgáló közlemények

Idegen nyelvű tudományos közlemény(ek) külföldi folyóiratban (3)

1. **Hegedüs, L.**, Nagy, B.: On periodic properties of circular words.
   Discret. Math. 339 (3), 1189-1197, 2016. ISSN: 0012-365X.
   DOI: http://dx.doi.org/10.1016/j.disc.2015.10.043
   IF:0.557 (2014)

2. **Hegedüs, L.**, Nagy, B., Eğecioğlu, Ö.: Stateless Multicounter 5'->3' Watson-Crick automata: The deterministic case.
   Nat Comput. 11 (3), 361-368, 2012. ISSN: 1567-7818.
   DOI: http://dx.doi.org/10.1007/s11047-011-9290-9
   IF:0.683

3. Nagy, B., **Hegedüs, L.**, Eğecioğlu, Ö.: Hierarchies of Stateless Multicounter 5’ -> 3’ Watson-Crick Automata Languages.
   Fundam. Inform. 110, 1-13, 2011. ISSN: 0169-2968.
   DOI: http://dx.doi.org/10.3233/FI-2011-553
   IF:0.365

Idegen nyelvű konferencia közlemény(ek) (5)

4. **Hegedüs, L.**, Nagy, B.: Representations of Circular Words.
   Electron. Proc. Theor. Comput. Sci. 151, 261-270, 2014. EISSN: 2075-2180.
   DOI: http://dx.doi.org/10.4204/EPTCS.151.18

5. **Hegedüs, L.**, Nagy, B.: On String Reading Stateless Multicounter 5'->3' Watson-Crick Automata.
In: Unconventional Computation and Natural Computation 12th International Conference, UCNC 2013, Milan, Italy, July 1-5, 2013. Proceedings. Ed.: Giancarlo Mauri, Alberto Dennunzio, Luca Manzoni, Antonio E. Porreca, Springer, Berlin, Heidelberg, 257-258, 2013.
DOI: http://dx.doi.org/10.1007/978-3-642-39074-6_29

6. **Hegedüs, L.**, Nagy, B.: Periodicity of Circular Words.
TUCS Lecture Notes. 20, 45-56, 2013. ISSN: 1797-8823.

7. Nagy, B., **Hegedüs, L.**, Eğecioğlu, Ö.: Hierarchy Results On Stateless Multicounter 5' -> 3' Watson-Crick Automata.
In: Advances in Computational Intelligence : 11th International Work-Conference on Artificial Neural Networks, IWANN 2011, Torremolinos-Málaga, Spain, June 8-10, 2011, Proceedings, Part I. Ed.: Joan Cabestany, Ignacio Rojas, Gonzalo Joya, Springer, Berlin, Heidelberg, 465-472, 2011. ISBN: 9783642215001
DOI: http://dx.doi.org/10.1007/978-3-642-21501-8_58

8. Eğecioğlu, Ö., **Hegedüs, L.**, Nagy, B.: Stateless Multicounter 5'->3' Watson-Crick Automata.
In: Bio-Inspired Computing: Theories and Applications. Ed.: Atulya K. Nagar et al, IEEE, Liverpool, 1599-1606, 2010.

További Közlemények

Idegen nyelvű konferencia közlemény(ek) (4)

9. Halász, V., **Hegedüs, L.**, Hornyák, I., Nagy, B.: Solving application oriented graph theoretical problems with DNA computing.
In: Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012) vol 1. Ed.: Jagdish Chand Bansal, Pramod Kumar Singh, Kusum Deep, Millie Pant, Atulya K. Nagar, Springer, India, 75-85, 2013. ISBN: 9788132210375(Print)

10. **Hegedüs, L.**, Nagy, B.: Periodicity of circular words.
In: Conference of PhD Students in Computer Science June 28 - June 30, 2012 Szeged, Hungary : Volume of extended abstracts. Szegedi Tudományegyetem, Szeged, 22, 2012.

11. **Hegedüs, L.**: On Some Subsystems of Interval-Valued Logic.
In: Conference of PhD Students in Computer Science June 28 - June 30, 2012 Szeged, Hungary : Volume of extended abstracts. Szegedi Tudományegyetem, Szeged, 21, 2012.

12. Halász, V., **Hegedüs, L.**, Hornyák, I., Nagy, B.: Solving application oriented graph theoretical problems with DNA computing.
In: How the world computes : Turing Centenary Conference And 8th Conference On Computability In Europe, CIE 2012, Cambridge, UK, June 18-23, 2012 : abstracts of informal presentations. Computer Laboratory, University of Cambridge, Cambridge, 105, 2012.

A közlő folyóiratok összesített impakt faktora: 1,605
A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre): 1,605

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudománymetriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2016.04.07.

Registry number: DEENK/87/2016.PL
Subject: Ph.D. List of Publications

Candidate: László Hegedüs
Neptun ID: YU4N37
Doctoral School: Doctoral School of Informatics
MTMT ID: 10037392

List of publications related to the dissertation

Foreign language scientific article(s) in international journal(s) (3)

1. **Hegedüs, L.**, Nagy, B.: On periodic properties of circular words.
   Discret. Math. 339 (3), 1189-1197, 2016. ISSN: 0012-365X.
   DOI: http://dx.doi.org/10.1016/j.disc.2015.10.043
   IF:0.557 (2014)

2. **Hegedüs, L.**, Nagy, B., Eğecioğlu, Ö.: Stateless Multicounter 5'->3' Watson-Crick automata: The deterministic case.
   Nat Comput. 11 (3), 361-368, 2012. ISSN: 1567-7818.
   DOI: http://dx.doi.org/10.1007/s11047-011-9290-9
   IF:0.683

3. Nagy, B., **Hegedüs, L.**, Eğecioğlu, Ö.: Hierarchies of Stateless Multicounter 5’ -> 3’ Watson-Crick Automata Languages.
   Fundam. Inform. 110, 1-13, 2011. ISSN: 0169-2968.
   DOI: http://dx.doi.org/10.3233/FI-2011-553
   IF:0.365

Foreign language conference proceeding(s) (5)

4. **Hegedüs, L.**, Nagy, B.: Representations of Circular Words.
   Electron. Proc. Theor. Comput. Sci. 151, 261-270, 2014. EISSN: 2075-2180.
   DOI: http://dx.doi.org/10.4204/EPTCS.151.18

5. **Hegedüs, L.**, Nagy, B.: On String Reading Stateless Multicounter 5'->3' Watson-Crick Automata. In: Unconventional Computation and Natural Computation 12th International Conference, UCNC 2013, Milan, Italy, July 1-5, 2013. Proceedings. Ed.: Giancarlo Mauri, Alberto Dennunzio, Luca Manzoni, Antonio E. Porreca, Springer, Berlin, Heidelberg, 257-258, 2013. DOI: http://dx.doi.org/10.1007/978-3-642-39074-6_29

6. **Hegedüs, L.**, Nagy, B.: Periodicity of Circular Words. *TUCS Lecture Notes. 20*, 45-56, 2013. ISSN: 1797-8823.

7. Nagy, B., **Hegedüs, L.**, Eğecioğlu, Ö.: Hierarchy Results On Stateless Multicounter 5' -> 3' Watson-Crick Automata. In: Advances in Computational Intelligence : 11th International Work-Conference on Artificial Neural Networks, IWANN 2011, Torremolinos-Málaga, Spain, June 8-10, 2011, Proceedings, Part I. Ed.: Joan Cabestany, Ignacio Rojas, Gonzalo Joya, Springer, Berlin, Heidelberg, 465-472, 2011. ISBN: 9783642215001 DOI: http://dx.doi.org/10.1007/978-3-642-21501-8_58

8. Eğecioğlu, Ö., **Hegedüs, L.**, Nagy, B.: Stateless Multicounter 5'->3' Watson-Crick Automata. In: Bio-Inspired Computing: Theories and Applications. Ed.: Atulya K. Nagar et al, IEEE, Liverpool, 1599-1606, 2010.

List of other publications

Foreign language conference proceeding(s) (4)

9. Halász, V., **Hegedüs, L.**, Hornyák, I., Nagy, B.: Solving application oriented graph theoretical problems with DNA computing. In: Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012) vol 1. Ed.: Jagdish Chand Bansal, Pramod Kumar Singh, Kusum Deep, Millie Pant, Atulya K. Nagar, Springer, India, 75-85, 2013. ISBN: 9788132210375(Print)

10. **Hegedüs, L.**, Nagy, B.: Periodicity of circular words. In: Conference of PhD Students in Computer Science June 28 - June 30, 2012 Szeged, Hungary : Volume of extended abstracts. Szegedi Tudományegyetem, Szeged, 22, 2012.

11. **Hegedüs, L.**: On Some Subsystems of Interval-Valued Logic.
In: Conference of PhD Students in Computer Science June 28 - June 30, 2012 Szeged, Hungary : Volume of extended abstracts. Szegedi Tudományegyetem, Szeged, 21, 2012.

12. Halász, V., **Hegedüs, L.**, Hornyák, I., Nagy, B.: Solving application oriented graph theoretical problems with DNA computing.
In: How the world computes : Turing Centenary Conference And 8th Conference On Computability In Europe, CIE 2012, Cambridge, UK, June 18-23, 2012 : abstracts of informal presentations. Computer Laboratory, University of Cambridge, Cambridge, 105, 2012.

Total IF of journals (all publications): 1,605
Total IF of journals (publications related to the dissertation): 1,605

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

07 April, 2016