Palindromy pod lupou
Brlekova-Reutenauerova hypotéza

Ľubomíra Balková

Katedra matematiky FJFI ČVUT v Praze

MELA v Telči 30.9.2012

Palindromy v různých jazycích

- latina:
  Signate, signate, mere me tangis et angis.
  Roma, tibi subito motibus ibit amor.

- čeština:
  Bažantu padá za záda putna žab.
  Jelenovi pivo nelej.
  Kobyla má malý bok.
- angličtina:
  tattarrattat

  I was just beginning to yawn with nerves thinking he was trying to make a fool of me when I knew his **tattarrattat** at the door.

Program

Program

Kombinatorika na slovech

- **abeceda** $\mathcal{A} = \{0, 1, \ldots, k-1\}$
- **slovo** = konečná posloupnost písmen, např. 0112001
- $\mathcal{A}^*$ = množina konečných slov nad $\mathcal{A}$
- slovo v je **faktor (podslovo)** slova w, pokud $w = pvs$, např. 112 nebo 2001 (**sufix**) nebo 0 (**prefix**)
- nekonečné slovo $\mathbf{u}$ je posloupnost $u_0 u_1 u_2 \ldots$, kde $u_i \in \mathcal{A}$
- $\mathcal{A}^{\mathbb{N}}$ = množina nekonečných slov nad $\mathcal{A}$
- $\mathcal{L}_n(\mathbf{u})$ = množina faktorů délky n slova $\mathbf{u}$
- $\mathcal{C}_{\mathbf{u}}(n) = \#\mathcal{L}_n(\mathbf{u})$ **(faktorová) komplexita slova u**

Příklad

$\mathbf{u} = 001001001001001\ldots$

$\mathcal{L}_1(\mathbf{u}) = \{0, 1\}, \mathcal{L}_2(\mathbf{u}) = \{00, 01, 10\}, \mathcal{L}_3(\mathbf{u}) = \{001, 010, 100\}, \mathcal{L}_4(\mathbf{u}) = \{0010, 0100, 1001\}, \mathcal{L}_5(\mathbf{u}) = \{00100, 01001, 10010\}$

$\mathcal{C}_{\mathbf{u}}(1) = 2, \mathcal{C}_{\mathbf{u}}(n) = 3 \textit{ pro } n \geq 2$

Substituce

- $\varphi : \mathcal{A}^* \to \mathcal{A}^*$ nazveme **substitucí**, pokud
  1. $\varphi(wv) = \varphi(w)\varphi(v)$ ($\Rightarrow$ stačí definovat $\varphi(a)$ pro každé $a \in \mathcal{A}$)
  2. existuje $a \in \mathcal{A}$ tak, že $\varphi(a) = aw$, kde $w \neq \varepsilon$
  3. $\varphi(a) \neq \varepsilon$ pro každé $a \in \mathcal{A}$

  Lze rozšířit na nekonečná slova:

$$\varphi(u_0 u_1 u_2 \dots) := \varphi(u_0)\varphi(u_1)\varphi(u_2)\dots .$$

  Každá substituce má aspoň jeden pevný bod:
  $\mathbf{u} := \lim_{n\to+\infty} \varphi^n(a)$ splňuje $\mathbf{u} = \varphi(\mathbf{u})$.

Příklad: Fibonacciho slovo $\mathbf{u}_F$ a Fibonacciho substituce

$\mathbf{u}_F$ je pevný bod substituce $\varphi(0) = 01,\ \varphi(1) = 0$

$$\begin{aligned}
\varphi(0) &= 01 \\
\varphi^2(0) &= 010 \\
\varphi^3(0) &= 01001 \\
\varphi^4(0) &= 01001010 \\
\varphi^5(0) &= 0100101001001
\end{aligned}$$

$$\mathbf{u}_F = 0100101001001010010100100101001001\dots$$

- $\mathcal{A} = \{0, 1\}$
- $\mathcal{L}_3(\mathbf{u}_F) = \{010, 100, 001, 101\}$
- $\mathcal{C}_{\mathbf{u}_F}(3) = 4$

Program

Sturmovská slova

u se nazývá **posléze periodické**, pokud $\mathbf{u} = wvvvvvvv\ldots$, jinak se nazývá **aperiodické**.

Věta (Hedlund, Morse, 1938)

1. **u** *je aperiodické* $\Leftrightarrow \mathcal{C}_{\mathbf{u}}(n) \geq n+1$ *pro každé* $n \in \mathbb{N}$.
2. *Posléze periodická slova mají omezenou komplexitu.*

Definice

Nechť $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ *splňuje* $\mathcal{C}_{\mathbf{u}}(n) = n+1$ *pro každé* $n \in \mathbb{N}$. *Pak* **u** *nazveme* **sturmovské**. *(Nutně* $\mathcal{A} = \{0,1\}$.*)*

Příklad

Fibonacciho slovo je sturmovské.

Palindromická komplexita

$$\mathcal{P}_{\mathbf{u}}(n) := \#\{w \in \mathcal{L}_n(\mathbf{u}) \mid w \text{ je palindrom}\}$$

Věta (Droubay, Pirillo, 1999)

Následující výroky jsou ekvivalentní:

1. **u** *je sturmovské.*
2. **u** *obsahuje jediný palindrom každé sudé délky a právě 2 palindromy každé liché délky.*

Příklad: Fibonacciho slovo $\mathbf{u}_F$ - pokračování

$\mathbf{u}_F$ je pevný bod substituce $\varphi(0) = 01,\ \varphi(1) = 0$

$$\mathbf{u}_F = 0100101001001010010100100101001001\ldots$$

- palindromy délky $3 = \{010, 101\}$
- $\mathcal{P}_{\mathbf{u}}(3) = 2$
- palindromy délky $4 = \{1001\}$
- $\mathcal{P}_{\mathbf{u}}(4) = 1$

Program

Vztah palindromické a faktorové komplexity

u má jazyk **uzavřený na reverzi**, pokud s každým faktorem obsahuje i jeho zrcadlový obraz, např. 0012 je faktor **u** $\Rightarrow$ 2100 je faktor **u**

Věta (Baláži, Masáková, Pelantová, 2007)

Nechť **u** *má jazyk uzavřený na reverzi. Pak*

$$\mathcal{P}_{\mathbf{u}}(n+1)+\mathcal{P}_{\mathbf{u}}(n)\leq \mathcal{C}_{\mathbf{u}}(n+1)-\mathcal{C}_{\mathbf{u}}(n)+2 \quad \textit{pro každé } n\in\mathbb{N}.$$

Definice

Slova, pro která nastává v předchozím vztahu rovnost, nazýváme **bohatá na palindromy**.

Příklad

Sturmovská slova jsou bohatá na palindromy:

- $\mathcal{P}_{\mathbf{u}}(n+1)+\mathcal{P}_{\mathbf{u}}(n)=3$
- $\mathcal{C}_{\mathbf{u}}(n+1)-\mathcal{C}_{\mathbf{u}}(n)+2=(n+2)-(n+1)+2=3$

Defekt

Věta (Droubay, Justin, Pirillo, 2001)

Každé konečné slovo w obsahuje maximálně $|w|+1$ palindromů (včetně prázdného slova).

Definice (Brlek, Hamel, Nivat, Reutenauer, 2004)

Nechť w je konečné slovo, pak **defektem** *slova w nazveme*

$$D(w) = |w| + 1 - \textit{počet palindromů obsažených ve } w.$$

Nechť $\mathbf{u}$ *je nekonečné slovo, pak* **defektem** *slova* $\mathbf{u}$ *nazveme*

$$D(\mathbf{u}) = \sup\{D(w) \;\big|\; w \textit{ je prefix slova } \mathbf{u}\}.$$

Defekt a slova bohatá na palindromy

Věta (Bucci, De Luca, Glen, Zamboni, 2009)

Nechť $\mathbf{u}$ *je nekonečné slovo s jazykem uzavřeným na reverzi. Pak* $D(\mathbf{u}) = 0$ *právě tehdy, když* $\mathbf{u}$ *je bohaté na palindromy, tj.* $\mathcal{P}_{\mathbf{u}}(n+1) + \mathcal{P}_{\mathbf{u}}(n) = \mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 2 \quad$ *pro každé* $n \in \mathbb{N}$.

Příklad

Uvažujme prefix $w = 01001$ *Fibonacciho slova.*
w *obsahuje 6 palindromů:* $\varepsilon, 0, 1, 00, 010, 1001,$ *tedy* $D(w) = 0$.
Sturmovská slova mají nulový defekt.

Defekt a slova bohatá na palindromy

Věta (Bucci, De Luca, Glen, Zamboni, 2009)

Nechť $\mathbf{u}$ *je nekonečné slovo s jazykem uzavřeným na reverzi. Pak* $D(\mathbf{u}) = 0$ *právě tehdy, když* $\mathbf{u}$ *je bohaté na palindromy, tj.*
$\mathcal{P}_{\mathbf{u}}(n+1) + \mathcal{P}_{\mathbf{u}}(n) = \mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 2 \quad$ *pro každé* $n \in \mathbb{N}$.

Příklad

Uvažujme prefix $w = 011010011$ *Thue-Morseova slova*
$\mathbf{u}_{TM} = \varphi(\mathbf{u}_{TM}) = 0110100110010110100101 10\ldots$, *kde*
$\varphi(0) = 01,\ \varphi(1) = 10$.
w obsahuje 9 palindromů: $\varepsilon, 0, 1, 00, 11, 010, 101, 0110, 1001$, *tedy*
$D(w) = 1$ *a* $D(\mathbf{u}_{TM}) \neq 0$.
Thue-Morseovo slovo není bohaté na palindromy.

Program

Brlekova-Reutenauerova hypotéza

Označme

$$T_{\mathbf{u}}(n) = \mathcal{C}_{\mathbf{u}}(n+1) - \mathcal{C}_{\mathbf{u}}(n) + 2 - \mathcal{P}_{\mathbf{u}}(n) - \mathcal{P}_{\mathbf{u}}(n+1).$$

Hypotéza (Brlek, Reutenauer, 2011)

Nechť $\mathbf{u}$ *je nekonečné slovo s jazykem uzavřeným na reverzi. Pak*

$$\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) = 2D(\mathbf{u}).$$

Ve svém článku dokázali platnost pro

- periodická slova,
- konečná slova,
- slova bohatá na palindromy.

Brlekova-Reutenauerova hypotéza - částečný důkaz

Věta (Balková, Pelantová, Starosta, 2011)

Nechť $\mathbf{u}$ *je stejnoměrně rekurentní s jazykem uzavřeným na reverzi. Pak*

$$\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) = 2D(\mathbf{u}).$$

V důkazu využito lemma:

Lemma

Nechť $\mathbf{u}$ *je nekonečné stejnoměrně rekurentní slovo s jazykem uzavřeným na reverzi a* $D(\mathbf{u}) < +\infty$. *Pak je množina* $\{w | ww \in \mathcal{L}(\mathbf{u})\}$ *nekonečná.*

Program

Důkaz Brlekovy-Reutenauerovy hypotézy

Věta (Balková, Pelantová, Starosta, 2012)

Brlekova-Reutenauerova hypotéza platí, tj. $\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) = 2D(\mathbf{u})$ *pro každé nekonečné slovo* $\mathbf{u}$ *s jazykem uzavřeným na reverzi.*

Důkaz.

2 kroky:

1. Pokud $D(\mathbf{u}) < \infty$ a $\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) < \infty$, pak $\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) = 2D(\mathbf{u})$.
2. $D(\mathbf{u}) < \infty$ právě tehdy, když $\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) < \infty$.

□

1. krok důkazu

Věta (1. krok)

Pokud $\mathbf{u}$ *je nekonečné slovo s jazykem uzavřeným na reverzi a* $D(\mathbf{u}) < \infty$ *a* $\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) < \infty$, *pak* $\sum_{n=0}^{\infty} T_{\mathbf{u}}(n) = 2D(\mathbf{u})$.

Důkaz.

Najdi H takové, že

1. $D(\mathbf{u}) = D(v)$ pro každý prefix v slova $\mathbf{u}$ délky $\geq H - 1$,
2. $T_{\mathbf{u}}(n) = 0$ pro každé $n \geq H$.

Uvažuj libovolný prefix p slova $\mathbf{u}$ obsahující všechny faktory délky $\leq H$.
Pak $2D(\mathbf{u}) = 2D(p) = \sum_{n=0}^{|p|} T_p(n) = \sum_{n=0}^{H-1} T_p(n) + \sum_{n=H}^{|p|} T_p(n) = \sum_{n=0}^{H-1} T_{\mathbf{u}}(n) = \sum_{n=0}^{\infty} T_{\mathbf{u}}(n)$.

□

Otevřené problémy

1. Existuje primitivní injektivní substituce φ, jejíž pevný bod $\mathbf{u} = \varphi(\mathbf{u})$ splňuje
$$0 < D(\mathbf{u}) < +\infty?$$
2. Existuje substituce φ nad abecedou $\mathcal{A}$, $\#\mathcal{A} > 2$, jejíž pevný bod $\mathbf{u} = \varphi(\mathbf{u})$ má jazyk uzavřený na reverzi a obsahuje konečně mnoho palindromů?
3. Charakterizujte slova “chudá na palindromy”.

Děkuji za pozornost

Ráček

Nedávejte drahé hračky
děcku **v žádném případě!**
Auta hází do zatáčky,
lego válí v zahradě.

Mašinky — ty rozebere.
Elefanta — rozmlátí.
Rodiče pak **trochu** berou
infarkty se závratí.

Kupte raděj míče, puky,
luky, kola horská.
Do přírody žeňte kluky —
Radí vám **muž z Norska**.

Nápověda: 2-5-6-4-3-5.

Řešení

Nedávejte drahé hračky
děcku **v žádném případě!**
Auta hází do zatáčky,
lego válí v zahradě.

Mašinky — ty rozebere.
Elefanta — rozmlátí.
Rodiče pak **trochu** berou
infarkty se závratí.

Kupte raděj míče, puky,
luky, kola horská.
Do přírody žeňte kluky —
Radí vám **muž z Norska**.

Nápověda: 2-5-6-4-3-5.
Řešení: **Ne slona poláme málo pan Olsen.**