Non-repetitive words: ages and essences

Wojciech Węgrzynek

na podstawie pracy James D. Currie

17 czerwca 2021

Plan prezentacji

- Definicje i przykłady
- Zformułowanie głównego wyniku pracy
- Opracowanie dowodów ciekawszych części

Definicje i przykłady

Przypomnienie:

- **Alfabet** - pewien skończony zbiór.
- **Litera** - element alfabetu.
- **Słowo skończone**- ciąg skończony liter.
- **Konkatenacja** słów skończonych u i v - słowo powstałe po przeczytaniu najpierw u później v, oznaczamy uv
- **Słowo puste**, oznaczamy przez ϵ.
- A^* - zbiór wszystkich słów skończonych nad alfabetem A.
- **Język** - podzbiór A^*.
- **Podsłowo** - $u \in A^*$ jest podsłowem $v \in A^*$ jeśli v możemy zapisać $v = xuy$ dla pewnych $x, y \in A^*$.
- **Prefiks, sufiks** - jw. tylko odpowiednio $x = \epsilon$ i $y = \epsilon$.
- Podsłowo, prefiks lub sufiks, są **właściwe** - jeśli nie są równy całemu słowu.

Definicje i przykłady (cont.)

Słowa nieskończone:

- **Słowo nieskończone** - nieskończony ciąg elementów alfabetu.
- A^ω - zbiór wszystkich słów nieskończonych na alfabecie A.
- **Język** - podzbiór A^ω.
- **Konkatenacja** uv - jak na słowach skończonych tylko dopuszczamy, że v jest nieskończone.
- Podsłowa, prefiksy, sufiksy - tak samo jak dla słów skończonych.

Definicje i przykłady (cont.)

Słowa nieskończone c.d.:

- **Morfizm** - funkcja $h : A^* \to B^*$ taka że dla każdych dwóch $u, v \in A^*$, $h(u)h(v) = h(uv)$. Mówiąc inaczej, h to funkcja zdefiniowana przez jej wartość na literach.
- **Granicą** w nieskończoności ciągu słów $w_1, w_2, w_3 \cdots$ takiego, że w_i jest właściwym prefiksem w_{i+1} będzie słowo nieskończone, które ma każde w_i jako prefiks.
- Dla litery $a \in A$ i morfizmu $h : A^* \to A^*$, jeśli a jest prefiksem właściwym $h(a)$ (wtedy granica istnieje), definiujemy $h^\omega(a) = \lim_{n\to\infty} h^n(a)$.

Definicje i przykłady (cont.)

Niech morfizm $h : \{0,1\}^* \to \{0,1\}^*$ będzie dany przez:

$$h(0) = 01$$

$$h(1) = 10$$

Wówczas:

$$h^1(0) = 01$$

$$h^2(0) = 0110$$

$$h^3(0) = 01101001$$

A więc:

$$h^\omega(0) = 0110100110010110\cdots$$

$h^\omega(0)$ jest znane jako słowo Thuego-Morse'a.

Definicje i przykłady (cont.)

Unikanie powtórzeń:

- r-potęga czy r-powtórzenie słowa u - słowo v powstałe po r-krotnej konkatenacji słowa u, oznaczamy $v = u^r$ (na potrzeby prezentacji $r \in \mathbb{N}_+$). Np. $(ab)(ab)(ab) = (ab)^3$.
- 2-powtórzenia nazywamy czasem kwadratami, a 3-powtórzenia sześcianami.
- Słowo jest r-wolne jeśli nie zawiera niepustego r-powtórzenia jako podsłowa.

Definicje i przykłady (cont.)

Wieki i esencje:

- Wiekiem słowa nieskończonego w nazwiemy zbiór wszystkich skończonych podsłów w.
- Esencją słowa nieskończonego w nazwiemy zbiór wszystkich skończonych podsłów w występujących w w nieskończenie wiele razy.

Rozważane języki

Przyjmujemy alfabet trzyelementowy $A = \{1, 2, 3\}$, wówczas

- L niech będzie zbiorem wszystkich 2-wolnych słów nieskończonych nad A.
- Niech $\{w_1, w_2, \cdots w_k\}$, wówczas $L_{w_1, w_2 \cdots w_k}$ to będzie podjęzyk L słów, które nie zawierają żadnego w_i jako podsłowa.

Zajmować będziemy się L, L_{121}, $L_{121,323}$ oraz $L_{121,212}$

Główny rezultat

equivalence classes with respect to	L	L_{121}	$L_{121,323}$	$L_{121,212}$
final segments	2^ω	2^ω	2^ω	2^ω
age	2^ω	2^ω	2^ω	2^ω
essence	2^ω	2^ω	1	1

Rysunek: Tabelka przedstawiająca wynik, Źródło: omawiana praca

3 propozycje

Propozycja 1

Jest nieprzeliczalnie wiele słów w $L_{121,323}$.

Propozycja 2

Jest nieprzeliczalnie wiele słów w $L_{121,323}$ *takich, że każde dwa mają różne wieki.*

Propozycja 3

Każde dwa słowa z $L_{121,323}$ *mają tą samą esencję*

Dowody Propozycji 1 i 2

Zdefiniujmy morfizmy $f_0, g_0 : A^* \to A^*$:

$$f_0(1) = 123$$
$$f_0(2) = 13$$
$$f_0(3) = 2$$

$$g_0(1) = 2$$
$$g_0(2) = 31$$
$$g_0(3) = 321$$

Dowody Propozycji 1 i 2 (cont.)

Najpierw 3 techniczne lematy ((prawie) bez dowodów):

Lemat 1

Niech $v \in A^*$ będzie 2-wolnym słowem nie zawierającym 121 ani 323 jako podsłowa, a $\phi \in \{f_0, g_0\}$. Wówczas $\phi(v)$ jest wolnym od kwadratów słowem nie zawierającym 121 ani 323 jako podsłowa. Zachodzi również $f_0^\omega(1), g_0^\omega(3) \in L_{121,323}$.

Lemat 2

Niech $v \in A^*$ będzie niepustym, 2-wolnym słowem nie zawierającym 121 ani 323. Niech $w = 323v$. $f_0(w)$ jest wówczas 2-wolne słowem nie zawierającym 323 ani 121

Lemat 3

Niech $w = u323v$ dla $u, v \in A^*$ niepustych, wówczas $f_0(w)$ ma kwadrat.

Dowód Lematu 3.

Jeżeli w ma kwadrat to $f_0(w)$ ma kwadrat. Załóżmy, że w jest 2-wolne, wówczas 323 musi być otoczone 1-ami, ale wówczas w zawiera 13231, a $f_0(13231) = 1232132123$, a więc zawiera 321321. $\square$

Dowód Propozycji 1

- $f_0^2(v)$ zaczyna się 1 (jeżeli v jest niepuste)
- Zdefiniujmy morfizmy $\mu_0 = f_0^3$ i $\mu_1 = f_0^2 g_0$
- Zdefiniujmy funkcję $\phi : \{0,1\}^* \to A^*$ rekurencyjnie:

$$\begin{aligned} \phi(\epsilon) &= 1 \\ \phi(0v) &= \mu_0(\phi(v)) \\ \phi(1v) &= \mu_1(\phi(v)) \end{aligned}$$

- Z Lematu 1. (prawdziwy również dla g_0) mamy, że $\phi(v)$ jest 2-wolnym słowem nie zawierającym 121 ani 323.
- ϕ zachowuje relację bycia prefiksem, tj. jeśli u jest prefiksem v to $\phi(u)$ jest prefiksem $\phi(v)$.
  Zauważmy, że jeśli $v = ux$, to

$$\begin{aligned} \phi(v) &= \mu_{u_1}(\mu_{u_2}(\mu_{u_3}(\cdots \mu_{u_m}(\phi(x)) \cdots))) \\ &= \mu_{u_1}(\mu_{u_2}(\mu_{u_3}(\cdots \mu_{u_m}(1x') \cdots))) \\ &= \phi(u)x'' \end{aligned}$$

Dowód Propozycji 1 (cont.)

- Skoro tak, możemy dla $v = v_1 v_2 \cdots v_n \cdots$ słowa nieskończonego, gdzie $v_i \in \{0, 1\}$ zdefiniować $\phi(v) = \lim_{n \to \infty} \phi(v_1 v_2 \cdots v_n)$.
- Każde dwa takie $\phi(v)$ są różne jeśli v są różne, a więc mamy nieprzeliczalnie wiele różnych słów w $L_{121,323}$.

Dowód Propozycji 2

- Zdefiniujmy $g_1 : A^* \to A^*$

$$g_1(w) = \begin{cases} 323w' \text{ dla w=123w'} \\ w \text{ wpp.} \end{cases}$$

- Definiujemy $h_0 = f_0^5$ i $h_1 = f_0^2 g_1 f_0^3$
- Definiujemy $\phi : \{0,1\}^* \to A^*$ tak jak poprzednio tylko za pomocą h_0 i h_1
- Podobnie jak poprzednio mamy nieprzeliczalnie wiele słów z $L_{121,323}$
- Chcemy jeszcze dodatkowo powiedzieć, że mają one różne wieki.
- Weźmy $u, v \in \{0,1\}^\omega$, $u \neq v$. Załóżmy że $u = u_1 u_2 \cdots u_k 0x$, a $v = u_1 u_2 \cdots u_k 1y$, dla $u_i \in \{0,1\}$ i weźmy $\psi = h_{u_1} h_{u_2} \cdots h_{u_k}$. A więc $\phi(u) = \psi(\phi(0x))$ i $\phi(v) = \psi(\phi(1y))$. Twierdzimy, że $\psi(f_0^2(323))$ jest podsłowem $\phi(v)$ ale nie $\phi(u)$.

Dowód Propozycji 2 (cont.)

- Okazuje się, że jeśli $\psi(f_0^2(v))$ zawiera $\psi(f_0^2(323))$ to v zawiera 323.
- Przypomnijmy, że $\phi(u) = \psi(f_0^2(f_0^3(\phi(x))))$ z Lematu 3 mamy, że $f_0^3(\phi(x))$ nie zawiera 323 „w środku"(inaczej $\phi(u)$ miałoby kwadrat), zaczyna się 1 więc nie zawiera 323 na początku, a więc nie zawiera 323, a więc $\phi(u)$ nie zawiera $\psi(f_0^2(323))$.
- Z kolei $\phi(v) = \psi(f_0^2(g_1(f_0^3(\phi(y)))))$ na pewno zawiera $\psi(f_0^2(323))$, gdyż $f_0^2(\phi(y))$ zaczyna się 1, więc $f_0^3(\phi(y))$ zaczyna się 123, a więc $g_1(f_0^3(\phi(y)))$ zaczyna się 323.

Dowód Propozycji 3

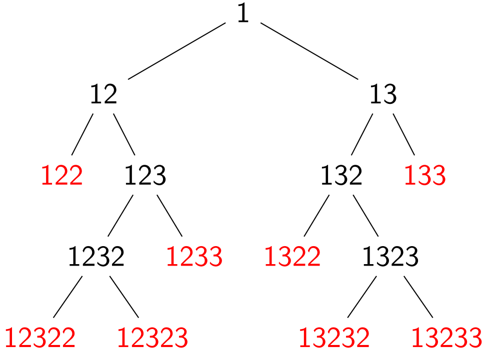
Okazuje się, że Propozycję 3 można sprowadzić do poniższego lematu.

Lemat 4

Niech $n \in \mathbb{N}$, jeśli $v \in L_{121,323}$ to istnieje sufiks v formy $f_0^n(w)$ gdzie $w \in L_{121,323}$

Dowód Propozycji 3 - dowód Lematu 4

- Zapiszmy v jako $v = 1b_1 1b_2 1b_3 \cdots 1b_n 1 \cdots$ gdzie b_i nie zawiera 1.
- Potencjalne 2-wolne $1b_i$ to 1, 12, 13, 123, 132, 1232, 1323.



Dowód Propozycji 3 - dowód Lematu 4 (cont.)

- v nie może zawierać 11, 121, ani 1323, a więc zostają $A = 13$, $B = 123$, $C = 132$, $D = 1232$.
- $BD = (123)(123)2$, $BA1 = 12(31)(31)$ mają kwadraty, więc B musi występować przed C.
- $AC = (13)(13)2$ i $DC1 = 12(321)(321)$ mają kwadraty, więc C musi występować po B.
- A więc mamy bloki A, BC i D.

Dowód Propozycji 3 - dowód Lematu 4 (cont.)

- Ale $A = f_0^2(3)$, $BC = f_0^2(1)$, a $D = f_0^2(2)$.
- A więc $v = f_0^2(w)$ dla pewnego w.
- $f_0(w)$ nie może zawierać 121 ani 323 (2 występuje albo jako część $f_0(1) = 123$, albo jako jako $f_0(3) = 2$, a więc następna litera to 1 albo 2, a poprzednia 2 albo 3).
- $f_0(w)$ musi być 2-wolne, w przeciwnym przypadku v nie byłoby.
- Indukcja daje Lemat 4.