Anežka Kasalová

Konstrukce slov s konstantními mezerami

Rozhledy matematicko-fyzikální, Vol. 97 (2022), No. 3, 1–12

Persistent URL: `http://dml.cz/dmlcz/151277`

Konstrukce slov s konstantními mezerami

Anežka Kasalová, Malostranské gymnázium, Praha

Abstrakt. Představíme si slova s konstantními mezerami a zaměříme se na jejich konstrukci s využitím počítače. Začneme terminologií a základními vlastnostmi slov s konstantními mezerami. Poté popíšeme pseudokód programu pro generování takových slov a uvedeme odkaz na jeho implementaci v Pythonu. Budeme se věnovat i optimalizacím programu a nakonec si představíme nové výzvy a problémy spojené s programem.

1. Úvod

Slova s konstantními mezerami, jimž se věnuje tento článek, jsou úzce spjatá s disciplínou zvanou kombinatorika na slovech. Základní pojmy vysvětlíme níže, pokud tedy nějaký termín z úvodu nebude jasný, laskavý čtenář si jej vyhledá dále v textu. V kombinatorice na slovech jsou velmi aktuálním tématem vlastnosti balancovaných slov, viz články z poslední doby [2, 4]. Nekonečné slovo nad abecedou $\mathcal{A}$ se nazývá balancované, pokud pro každou dvojici jeho faktorů u a v stejné délky platí, že pro libovolné písmeno z abecedy $\mathcal{A}$ se počet výskytů tohoto písmene ve slovech u a v liší nejvýše o jedna. Nad binární abecedou $\{\mathtt{a}, \mathtt{b}\}$ jsou aperiodická balancovaná slova právě dobře známá slova sturmovská [1]. Pro vícepísmenné abecedy platí, že aperiodická balancovaná slova lze konstruovat tak, že se výskyty písmene `a` ve sturmovském slově postupně přepisují pomocí nějakého slova s konstantními mezerami a výskyty písmene `b` pomocí jiného slova s konstantními mezerami nad disjunktními abecedami.

Příklad 1. Nejznámějším sturmovským slovem je Fibonacciho slovo, které získáme, když začneme s písmenem `a` a postupně aplikujeme přepisovací pravidla $\mathtt{a} \to \mathtt{ab}$ a $\mathtt{b} \to \mathtt{a}$ písmeno po písmenu. Dostáváme tak delší a delší prefixy nekonečného Fibonacciho slova:

```
a
ab
aba
abaab
abaababa
abaababaabaab
abaababaabaababaababa
...
```

Přepišme nyní Fibonacciho slovo pomocí slov s konstantními mezerami $(01)^\omega$ a $(234235)^\omega$.

$$\texttt{abaababaabaababaababaabaababaab}\ldots$$
$$\texttt{0210314012013051021301401203105}\ldots$$

Čtenář může zkontrolovat, že faktory stejné délky v prefixu Fibonacciho slova i v prefixu získaného slova splňují podmínku balancovanosti.

2. Pojmy z kombinatoriky na slovech

V kombinatorice na slovech se studují vlastnosti konečných a nekonečných slov. Právě do této oblasti patří námi studovaná slova s konstantními mezerami. Pojďme se tedy nejprve podívat na základy kombinatoriky na slovech.

Abecedou $\mathcal{A}$ rozumíme konečnou množinu symbolů, říkáme jim *písmena*. *Slovem* w nazýváme konečnou posloupnost písmen. Jeho *délkou* rozumíme počet písmen, která obsahuje, a označíme ji $|w|$. Pod *nekonečným slovem* $\mathbf{u}$ nad abecedou $\mathcal{A}$ pak rozumíme nekonečnou posloupnost písmen z $\mathcal{A}$, v níž se každé písmeno z $\mathcal{A}$ skutečně vyskytuje, tj. $\mathbf{u} = u_1u_2u_3\ldots$, kde $u_i \in \mathcal{A}$. Nekonečné slovo $\mathbf{u}$ nazveme *periodické*, pokud existuje konečné slovo v takové, že $\mathbf{u} = vvvvvvv\cdots = v^\omega$. Symbolem ω tedy značíme nekonečné řetězení stejného slova. Nekonečné slovo se nazývá *aperiodické*, pokud není periodické, a to ani po useknutí libovolného konečného prefixu.[1)]

Nechť $\mathbf{u} = u_1u_2u_3\ldots$ je periodické nekonečné slovo nad abecedou $\mathcal{A} = \{0, 1, \ldots, d-1\}$. *Periodou* P slova $\mathbf{u}$ nazveme nejmenší $P \in \mathbb{N}$ takové, že $u_j = u_{j+P}$ pro každé $j \in \mathbb{N}$.[2)]

Říkáme, že nekonečné slovo $\mathbf{u}$ má *konstantní mezery*, pokud pro každé písmeno $i \in \mathcal{A}$ platí, že existuje konstanta $p_i \in \mathbb{N}$ tak, že mezery mezi po sobě jdoucími výskyty písmene i v $\mathbf{u}$ jsou rovny p_i. Takové p_i nazveme *periodou písmene* i. Je zřejmé, že slovo $\mathbf{u}$ s konstantními mezerami je periodické a že perioda každého písmene dělí periodu slova $\mathbf{u}$. Dále první výskyt písmene $i \in \mathcal{A}$ budeme značit n_i. Je snadné nahlédnout, že $n_i \le p_i$.

[1)] V textu jsou nekonečná slova značena tučně, konečná normálně.

[2)] Pojem perioda budeme rovněž používat pro slova, např. pro $\mathbf{u} = (0102)^\omega$ je periodou $P = 4$, ale také slovo 0102. Konkrétní význam bude vždy zřejmý z kontextu.

Příklad 2. Příkladem slova s konstantními mezerami nad abecedou $\{0,1,2\}$ je $\mathbf{u} = (0102)^\omega$. Perioda slova $\mathbf{u}$ je $P = 4$ a perioda písmene nula je $p_0 = 2$ a perioda písmene jedna a dva je rovna $p_1 = p_2 = 4$. První výskyty písmen jsou postupně $n_0 = 1, n_1 = 2, n_2 = 4$. Příkladem periodického slova, které nemá konstantní mezery, nad abecedou $\{0,1,2\}$ je $\mathbf{v} = (0122)^\omega$.

3. Vlastnosti slov s konstantními mezerami

Definice 1. Nekonečná slova s konstantními mezerami, která vzniknou jedno z druhého permutací abecedy nebo posunem nebo obojím, nazveme *ekvivalentní.*

Příklad 3. Slova $\mathbf{u} = (0102)^\omega$ a $\mathbf{v} = (2101)^\omega$ jsou ekvivalentní, protože $\mathbf{v}$ vzniklo z $\mathbf{u}$ posunem o jednu pozici a permutací $0 \to 1, 1 \to 2, 2 \to 0$.

Pozorování 1. Ověřme, že jde o ekvivalenci na množině nekonečných slov s konstantními mezerami nad abecedou $\mathcal{A}$.[3)]

1. Reflexivita: Zjevně $\mathbf{u}$ je ekvivalentní s $\mathbf{u}$, protože vzniklo identickou permutací písmen.
2. Symetrie: Pokud $\mathbf{v}$ je ekvivalentní s $\mathbf{u}$, pak $\mathbf{v}$ vzniklo z $\mathbf{u}$ posunem o k pozic (k může být i nula) a jistou permutací písmen (může být i identická). Jistě $\mathbf{u}$ a $\mathbf{v}$ mají stejnou délku periody P. Pak $\mathbf{u}$ vzniklo z $\mathbf{v}$ inverzní permutací písmen a posunem o $P - k$ pozic, tedy $\mathbf{v}$ je ekvivalentní s $\mathbf{u}$.
3. Tranzitivita: Pokud $\mathbf{u}$ je ekvivalentní s $\mathbf{v}$ a $\mathbf{v}$ je ekvivalentní s $\mathbf{w}$, pak $\mathbf{v}$ vzniklo z $\mathbf{u}$ posunem o k pozic a jistou permutací písmen π, dále $\mathbf{w}$ vzniklo z $\mathbf{v}$ posunem o ℓ pozic a jistou permutací písmen σ. Tudíž $\mathbf{w}$ vzniklo z $\mathbf{u}$ posunem o $k + \ell$ pozic a složenou permutací písmen $\sigma \circ \pi$ a platí tedy, že $\mathbf{w}$ je ekvivalentní s $\mathbf{u}$. Závěr zní, že relace je i tranzitivní.

Příklad 4. Jedinými slovy s konstantními mezerami s písmeny z abecedy $\{0,1\}$ jsou $\mathbf{u} = (01)^\omega$ a $\mathbf{v} = (10)^\omega$. Jsou ekvivalentní, protože $\mathbf{v}$ vzniklo z $\mathbf{u}$ posunem o jednu pozici nebo též permutací $0 \to 1$, $1 \to 0$. Slova s konstantními mezerami s písmeny z abecedy $\{0,1,2\}$ mají dvě

[3)] Ekvivalence E na neprázdné množině M je každá podmnožina $E \subset M \times M = \{(m,n) \mid m,n \in M\}$, která má vlastnosti reflexivity, symetrie a tranzitivity. Místo $(m,n) \in E$ píšeme $m \sim n$. Vzájemně ekvivalentní prvky tvoří tzv. třídu ekvivalence.

třídy ekvivalence. Jejich reprezentanty jsou $\mathbf{u} = (012)^\omega$ a $\mathbf{v} = (0102)^\omega$. Další slova s konstantními mezerami už z nich získáme jedině permutací abecedy, např. $(201)^\omega$, nebo posunem, např. $(1020)^\omega$, nebo obojím, např. $(0121)^\omega$.

V programu pro hledání všech slov s konstantními mezerami se nám budou velmi hodit následující tvrzení o periodách slov a písmen.

Lemma 1. *Délka periody nekonečného slova s konstantními mezerami je vždy rovna nejmenšímu společnému násobku period jednotlivých písmen.*

Důkaz. Nechť perioda slova je P a n je nsn všech period písmen. P je dělitelné všemi periodami písmen, proto n dělí P. Takže $P \geq n$.

Na druhou stranu, vybereme-li jakékoliv písmeno a posuneme se od něj o n, pak narazíme znovu na to samé písmeno. Neboli $u_i = u_{i+n}$ pro každé $i \in \mathbb{N}$. A jelikož P je nejmenší číslo s předchozí vlastností, je jasné, že $P \leq n$.

Z těchto dvou úvah pak plyne, že $n = P$.

Lemma 2. *Periody písmen ve slově s konstantními mezerami nad abecedou $\mathcal{A} = \{0, 1, \ldots, d-1\}$ splňují* [4]

$$\sum_{i=0}^{d-1} \frac{1}{p_i} = 1.$$

Důkaz. Pokud slovo $\mathbf{u}$ má konstantní mezery a periodu P a pokud písmeno i má periodu p_i, pak se i v prefixu $\mathbf{u}$ délky P vyskytuje $\frac{P}{p_i}$-krát. Odtud plyne, že

$$P = \sum_{i=0}^{d-1} \frac{P}{p_i},$$

což po vykrácení číslem P dává $\sum_{i=0}^{d-1} \frac{1}{p_i} = 1$.

Definice 2. Nechť n, n', p, p' jsou přirozená čísla splňující $n \leq p$ a $n' \leq p'$. Řekneme, že (n, p) a (n', p') jsou *v kolizi*, pokud existují $k, k' \in \mathbb{N}_0$ [5] taková, že

$$n + kp = n' + k'p'. \tag{1}$$

[4] Používáme zkrácený zápis součtu $\sum\limits_{i=0}^{d-1} \frac{1}{p_i} = \frac{1}{p_0} + \frac{1}{p_1} + \cdots + \frac{1}{p_{d-1}}$.

[5] Symbol $\mathbb{N}_0$ značí množinu obsahující přirozená čísla a nulu.

Lemma 3. *Páry (n,p) a (n',p') jsou v kolizi, právě když $D = \operatorname{nsd}(p,p')$* [6] *je dělitelem $n - n'$.*

Důkaz. Pokud jsou (n,p) a (n',p') v kolizi, podle (1) platí $n - n' = k'p' - kp$. Jelikož D dělí p i p', jistě D dělí $n - n'$. Pro důkaz opačné implikace využijeme, že když $D = \operatorname{nsd}(p,p')$, pak použitím Eukleidova algoritmu je možné najít $\hat{k}, \hat{\ell} \in \mathbb{Z}$ tak, že $\hat{k}p' = D + \hat{\ell}p$. Pak ale také umíme najít $\tilde{k}, \tilde{\ell} \in \mathbb{Z}$ tak, že $\tilde{k}p' = n - n' + \tilde{\ell}p$ (předchozí rovnost jsme vynásobili celým číslem $\frac{n-n'}{D}$). Odtud $n + \tilde{\ell}p = n' + \tilde{k}p'$. Pokud $\tilde{k}, \tilde{\ell}$ nejsou obě nezáporná, pak stačí přičíst na obě strany vhodný násobek pp' tak, aby $n + (\tilde{\ell} + sp')p = n' + (\tilde{k} + sp)p'$ splňovalo, že $k = (\tilde{\ell} + sp') \geq 0$ a $k' = (\tilde{k} + sp) \geq 0$. Dostali jsme $n + kp = n' + k'p'$ pro čísla $k, k' \in \mathbb{N}_0$, což podle (1) znamená, že (n,p) a (n',p') jsou v kolizi.

Z předchozích lemmat bezprostředně plyne stěžejní věta, která popisuje slova s konstantními mezerami a na níž založíme program pro generování slov s konstantními mezerami.

Věta 1. *Nechť je dána abeceda $\mathcal{A} = \{0, 1, \ldots, d-1\}$ a páry (n_0, p_0), (n_1, p_1), …, (n_{d-1}, p_{d-1}) takové, že $n_i \leq p_i$ pro každé $i \in \mathcal{A}$. Pak páry (n_0, p_0), (n_1, p_1), …, (n_{d-1}, p_{d-1}) dají vzniknout slovu s konstantními mezerami, právě když žádné dva páry (n_i, p_i) a (n_j, p_j) pro $i, j \in \mathcal{A}$, $i \neq j$, nejsou v kolizi a $\sum_{i=0}^{d-1} \frac{1}{p_i} = 1$.*

Důsledek 1. *V nekonečném slově s konstantními mezerami musí být všechny periody písmen po dvojicích soudělné.*

Důkaz. Pokud existuje dvojice nesoudělných period písmen p_i, p_j, tj. $D = \operatorname{nsd}(p_i, p_j) = 1$, pak podle lemmatu 3 jsou páry (n_i, p_i) a (n_j, p_j) v kolizi, což je dále podle věty 1 nepřípustné pro slovo s konstantními mezerami.

Důsledek 2. *Nekonečné slovo* $\mathbf{u}$ *s konstatními mezerami nad abecedou $\mathcal{A} = \{0, 1, \ldots, d-1\}$ může mít $\operatorname{nsd}(p_0, p_1, \ldots, p_{d-1}) = 1$, pouze když se v prvočíselném rozkladu periody p_i každého písmene $i \in \mathcal{A}$ nachází alespoň dvě různá prvočísla.*

Důkaz. Pro spor předpokládejme, že by prvočíselný rozklad p_i jednoho z písmen obsahoval mocninu pouze jednoho prvočísla. Protože $\operatorname{nsd}(p_0, p_1, \ldots, p_{d-1}) = 1$, jistě musí existovat p_j nesoudělné s p_i. Podle důsledku 1 toto nemůže nastat.

[6] $\operatorname{nsd}(p,p')$ značí největšího společného dělitele p a p'.

Lemma 4. *Nechť α je prvočíslo, které se vyskytuje jako dělitel některé z period písmen. Označme m maximální exponent takový, že α^m dělí některou z period písmen. Pak α^m dělí alespoň dvě periody p_i a p_j pro $i \neq j$, $i, j \in \mathcal{A} = \{0, 1, \ldots, d-1\}$.*

Důkaz. Jelikož $P = \mathrm{nsn}(p_0, p_1, \ldots, p_{d-1})$, číslo α^m dělí P, zatímco číslo α^{m+1} nedělí P. Protože $\sum_{i=0}^{d-1} \frac{1}{p_i} = 1$, platí vztah:

$$P = \sum_{i=0}^{d-1} \frac{P}{p_i}.$$

Pokud pouze jedna z period, bez újmy na obecnosti p_1, je dělitelná α^m, pak ve výše uvedené rovnici jsou všechny výrazy až na $\frac{P}{p_1}$ dělitelné α, což je spor.

Na závěr této kapitoly zmíníme metodu, pomocí které lze získávat slova s konstantními mezerami nad většími abecedami pomocí slov s konstantními mezerami nad menšími abecedami. Je to proplétáním (anglicky shuffling) konečného počtu slov s konstantními mezerami nad disjunktními abecedami.

Definice 3. Mějme k nekonečných slov. Jejich *proplétáním* rozumíme slovo, které vznikne postupným čtením prvních písmen z daných k slov, poté druhých písmen, třetích atd.

Pokud propleteme k slov s konstantními mezerami nad disjunktními abecedami, vznikne slovo s konstantními mezerami, v němž se perioda každého písmene oproti původnímu slovu k-krát zvětšila.

Příklad 5. Slovo s konstantními mezerami $\mathbf{u} = (013024013025)^\omega$ vzniklo proplétáním tří slov $(0)^\omega$, $(12)^\omega$, $(3435)^\omega$. Skutečně platí, že se periody proplétáním třikrát zvětšily:

- perioda písmene 0 ve slově $\mathbf{u}$ je rovna 3, zatímco $p_0 = 1$ ve slově $(0)^\omega$,
- perioda písmene 1 a 2 ve slově $\mathbf{u}$ je rovna 6, zatímco $p_1 = p_2 = 2$ ve slově $(12)^\omega$,
- perioda písmene 3 ve slově $\mathbf{u}$ je rovna 6, zatímco $p_3 = 2$ ve slově $(3435)^\omega$,
- perioda písmene 4 a 5 ve slově $\mathbf{u}$ je rovna 12, zatímco $p_4 = p_5 = 4$ ve slově $(3435)^\omega$.

Naopak také platí, že když k dělí periodu každého písmene ve slově s konstantními mezerami, pak toto slovo vzniklo proplétáním k slov s konstantními mezerami, tj. pokud $\mathbf{u} = u_1u_2u_3\ldots$, pak $\mathbf{u}$ je propletením slov:

$$u_1u_{1+k}u_{1+2k}u_{1+3k}\cdots$$
$$u_2u_{2+k}u_{2+2k}u_{2+3k}\cdots$$
$$\cdots$$
$$u_ku_{2k}u_{3k}u_{4k}\cdots$$

Příklad 6. Slovo $\mathbf{u} = (012304210324)^\omega$ splňuje $p_0 = p_2 = 4$ a $p_1 = p_3 = p_4 = 6$, tedy periody jsou sudé. Slovo $\mathbf{u}$ vzniklo propletením slov $(02)^\omega$ a $(134)^\omega$.

4. Program

Hlavním cílem článku je představit náš program pro generování slov s konstantními mezerami. Je dostupný i s popisem přes webovou službu GitHub [3].

Přistupme nejprve k popisu pseudokódu.

4.1. Pseudokód programu

Vstupem programu je počet písmen abecedy $d \geq 2$. (Pro $d = 1$ existuje jediné slovo s konstantními mezerami 0^ω.) Bez újmy na obecnosti předpokládáme, že písmeno 0 má nejmenší periodu p_0 a jeho první výskyt $n_0 = 1$. Výstupem programu jsou všechna nekonečná slova s konstantními mezerami (až na ekvivalenci) nad abecedou dané velikosti, přesněji řečeno, vypíše se tvar jejich period.

Přípravná funkce:

- Do proměnné d vlož uživatelem zadaný počet písmen a $\mathcal{A}$ polož rovno $\{0, 1, \ldots, d-1\}$.
- Vytvoř seznam Prefixy a vlož do něj slova

$$010, 0120, 01230, \ldots, 0123\ldots(d-1)0.$$

- Vytvoř prázdný seznam Periody.

Dokud není seznam Prefixy prázdný, aplikuj vždy na první slovo ze seznamu následující funkci.

Funkce přidání dalšího písmene:

Pokud slovo u ze seznamu Prefixy neobsahuje již každé písmeno alespoň dvakrát, proveď následující kroky:

- Najdi nejdelší slovo v s prefixem u, které je jednoznačně určeno parametry (n_i, p_i), kde i jsou opakující se písmena ve slově u. Pokračuj se slovem v.
- Pokud v neobsahuje všechna písmena z $\mathcal{A}$, na pozici $|v| + 1$ doplň nejmenší písmeno i z $\mathcal{A}$, které v neobsahuje. Nové slovo vi vlož na konec seznamu Prefixy.
- Uvažuj všechna písmena j, která se ve v vyskytují jednou a splňují:

  (a) $p_j = |v| + 1 - n_j \geq n_j$, tj. že po připsání j za slovo v je jeho perioda p_j větší nebo rovna prvnímu výskytu j,

  (b) $p_j \geq p_0$, tj. že perioda p_j je větší nebo rovna periodě p_0.

  Pro každé takové j zkontroluj, zda jemu odpovídající pár

  $$(n_j, p_j) = (n_j, |v| + 1 - n_j)$$

  není v kolizi s žádným párem (n_i, p_i), kde i jsou písmena obsažená ve v alespoň dvakrát.

  – Pokud nenastane žádná kolize a

  $$\sum_i \frac{1}{p_i} + \frac{1}{|v| + 1 - n_j} < 1,$$

  přidej nové slovo vj na konec seznamu Prefixy.

  – Pokud nenastane žádná kolize a

  $$\sum_i \frac{1}{p_i} + \frac{1}{|v| + 1 - n_j} = 1$$

  a slovo vj obsahuje všechna písmena z $\mathcal{A}$, do seznamu Periody vlož prefix vj délky $\mathrm{nsn}(p_0, p_1, \ldots, p_{d-1})$.
- Slovo u smaž ze seznamu Prefixy.

Funkce, která vyhodí ekvivalentní slova:

- V seznamu Periody seřaď slova podle délky.
- Pokud je jen jedno slovo dané délky, vytiskni ho.
- Slova, která mají stejnou délku, porovnej pomocí Funkce porovnání.

Funkce porovnání:

- Ze slov stejné délky jedno vytiskni.
- Procházej postupně následující slova stejné délky ze seznamu Periody a porovnávej je s vytištěnými slovy:
  - Pokud je slovo stejné jako některé již vytištěné, pak ho zahoď.
  - U porovnávaného slova přehoď první písmeno na konec, proveď permutaci písmen tak, aby první výskyty písmen splňovaly $n_0 < n_1 < \ldots < n_{d-1}$, a pokud je nyní stejné jako některé již vytištěné, tak ho zahoď. Takto pokračuj tolikrát, kolik je délka periody.
  - Pokud v žádném případě nedojde ke shodě, pak slovo vytiskni.

4.2. Tvoření slov podle pseudokódu programu

Pro čtenářovu lepší orientaci popišme konstrukci všech nekonečných slov s konstatními mezerami nad abecedou o třech písmenech podle programu.

- Seznam Prefixy na začátku obsahuje slova 010 a 0120.
- Začneme s prefixem 010. Podle Funkce přidání dalšího písmene přidáme do seznamu Prefixy slovo 0102. Dále vytvoříme slovo 0101, které ale splňuje $\sum_{i=0}^{1} \frac{1}{p_i} = 1$, a přitom neobsahuje tři písmena. Slovo 010 smažeme ze seznamu Prefixy.
- Poté pokračujeme prefixem 0120 – podle Funkce přidání dalšího písmene vytvoříme dva nové prefixy 01201 a 01202. Druhý z nich ale nesplní podmínku $n_2 \leq p_2$, protože $n_2 = 3$ a $p_2 = 2$. Do Prefixů přidáme proto pouze 01201 a naopak smažeme 0120.
- Prefix 0102 podle Funkce přidání dalšího písmene prodloužíme na 01020 a poté dostaneme 010201 a 010202, kde ovšem druhý prefix opět nesplní podmínku $n_2 \leq p_2$. Do Prefixů přidáme proto pouze 010201 a naopak smažeme 0102.
- Prefix 01201 doplníme podle Funkce přidání dalšího písmene na 012012 a do seznamu Periody vložíme slovo 012. Smažeme 01201 z Prefixů.
- Prefix 010201 povinně prodloužíme na 0102010. Následně doplníme na 01020102 a do seznamu Periody vložíme 0102. Smažeme 010201 ze seznamu Prefixy, čímž jej vyprázdníme.

- Jelikož máme od každé délky jedinou periodu, vytiskneme je.

  Závěr: Vytiskli jsme slova 012 a 0102.

4.3. Vylepšení programu při implementaci

- Ve skutečnosti program neukládá celý tvar prefixu. V pseudokódu je to uvedeno pro přehlednost. Program si pamatuje n_i a p_i u každého písmene.
- Rovnice $\sum_{i=0}^{d-1} \frac{1}{p_i} = 1$ je v programu upravena na

$$\sum_{i=0}^{d-1} \frac{1}{p_i} + 0{,}000\,000\,1 \geq 1,$$

  a to vzhledem k reprezentaci reálných čísel počítačem. (Počítač k výpočtu nepoužívá zlomky, ale jejich přibližný přepočet ve float aritmetice.)
- Do programu je přidána možnost volby mezi třemi výstupy. První zobrazí periody slov s konstantními mezerami a druhý nejprve celkovou délku periody a poté ke každému písmenu napíše jeho n_i a p_i. Třetí výstup vypíše možné délky period.
- Ve Funkci porovnání nahrazujeme jednotlivá písmena jejich číselnou periodou. A následné porovnávání provádíme nikoli posouváním o jedno písmeno, ale posouváním o nejmenší periodu písmene (což je perioda písmene 0).

4.4. Možné optimalizace programu vedoucí ke zvýšení rychlosti

Program, který máme implementovaný, na průměrném počítači najde slova s konstantními mezerami nad abecedou s nejvýše 12 písmeny. Jako nadějná optimalizace do budoucna vypadá využití proplétání. Nezískáme tímto způsobem sice všechna slova s konstantními mezerami (nelze získat slova, pro která je největší společný dělitel period písmen roven jedné), ale i tak by mohlo jít o výrazné urychlení.

5. Úlohy pro čtenáře

Úloha pro pozorné a vědomostí chtivé čtenáře: *Za pomoci teorie z článku vymyslete příklad nekonečného slova s konstantními mezerami tak, aby největší společný dělitel period písmen byl jedna.*

Poděkování

Závěrem bych chtěla poděkovat doc. Ing. Ľubomíře Dvořákové, Ph.D., za toto hezké výzkumné téma, kterému se věnuji i ve své práci SOČ. Děkuji také za milou pomoc při psaní článku. Poděkování patří také prof. Ing. Editě Pelantové, CSc., za užitečné návrhy ke zjednodušení článku i programu. Poděkuji ještě svému obětavému tatínkovi panu Štěpánu Kasalovi za pomoc s kódováním programu.

Literatura

[1] Balková, Ľ.: Nahlédnutí pod pokličku kombinatoriky na nekonečných slovech. *Pokroky matematiky, fyziky a astronomie*, 56 (2011), č. 1, s. 9–18.

[2] Dolce, F., Dvořáková, Ľ., Pelantová, E.: On balanced sequences and their asymptotic critical exponent. *Proceedings LATA 2021, LNCS*, 12638 (2021), s. 293–304.

[3] Kasalová, A.: *Programy pro hledání slov s konstantními mezerami.* `https://github.com/kasal/ak-soc`

[4] Rampersad, N., Shallit, J., Vandomme, É.: Critical exponents of infinite balanced words. *Theoret. Comput. Sci.*, 777 (2019), s. 454–463.

Příloha

1. výstup programu pro abecedu délky 4:

```
SEQ: [0, 1, 0, 2, 0, 1, 0, 3]
SEQ: [0, 1, 0, 2, 0, 3]
SEQ: [0, 1, 2, 0, 1, 3]
SEQ: [0, 1, 2, 3]
```

2. výstup programu pro abecedu délky 4:

```
seq 0, with period length 8
  n_i:  0  1  3  7
  p_i:  2  4  8  8
seq 1, with period length 6
  n_i:  0  1  3  5
  p_i:  2  6  6  6
```

```
seq 2, with period length 6
  n_i:  0  1  2  5
  p_i:  3  3  6  6
seq 3, with period length 4
  n_i:  0  1  2  3
  p_i:  4  4  4  4
```

3. výstup programu pro abecedu délky 4:

```
Possible period lengths for alphabet size 4:
 [4, 6, 8]
```

Dále uvedeme 2. typ výstupu pro další dvě velikosti abecedy.

slovo	délka periody	n_0, n_1, n_2, n_3, n_4	p_0, p_1, p_2, p_3, p_4
1.	16	$0, 1, 3, 7, 15$	$2, 4, 8, 16, 16$
2.	12	$0, 1, 3, 7, 11$	$2, 4, 12, 12, 12$
3.	12	$0, 1, 3, 5, 11$	$2, 6, 6, 12, 12$
4.	8	$0, 1, 3, 5, 7$	$2, 8, 8, 8, 8$
5.	12	$0, 1, 2, 5, 11$	$3, 3, 6, 12, 12$
6.	9	$0, 1, 2, 5, 8$	$3, 3, 9, 9, 9$
7.	6	$0, 1, 2, 4, 5$	$3, 6, 6, 6, 6$
8.	8	$0, 1, 2, 3, 7$	$4, 4, 4, 8, 8$
9.	12	$0, 1, 2, 3, 5$	$4, 6, 4, 6, 6$
10.	5	$0, 1, 2, 3, 4$	$5, 5, 5, 5, 5$

Tabulka 1: Výstup pro $\mathcal{A} = \{0, 1, 2, 3, 4\}$

slovo	délka periody	$n_0, n_1, n_2, n_3, n_4, n_5$	$p_0, p_1, p_2, p_3, p_4, p_5$
1.	32	$0, 1, 3, 7, 15, 31$	$2, 4, 8, 16, 32, 32$
2.	24	$0, 1, 3, 7, 15, 23$	$2, 4, 8, 24, 24, 24$
3.	24	$0, 1, 3, 7, 11, 23$	$2, 4, 12, 12, 24, 24$
4.	16	$0, 1, 3, 7, 11, 15$	$2, 4, 16, 16, 16, 16$
5.	24	$0, 1, 3, 5, 11, 23$	$2, 6, 6, 12, 24, 24$
6.	18	$0, 1, 3, 5, 11, 17$	$2, 6, 6, 18, 18, 18$
7.	12	$0, 1, 3, 5, 9, 11$	$2, 6, 12, 12, 12, 12$
8.	16	$0, 1, 3, 5, 7, 15$	$2, 8, 8, 8, 16, 16$
9.	24	$0, 1, 3, 5, 7, 11$	$2, 8, 12, 8, 12, 12$
10.	10	$0, 1, 3, 5, 7, 9$	$2, 10, 10, 10, 10, 10$
11.	24	$0, 1, 2, 5, 11, 23$	$3, 3, 6, 12, 24, 24$
12.	18	$0, 1, 2, 5, 11, 17$	$3, 3, 6, 18, 18, 18$
13.	18	$0, 1, 2, 5, 8, 17$	$3, 3, 9, 9, 18, 18$
14.	12	$0, 1, 2, 5, 8, 11$	$3, 3, 12, 12, 12, 12$
15.	12	$0, 1, 2, 4, 5, 11$	$3, 6, 6, 6, 12, 12$
16.	12	$0, 1, 2, 4, 5, 10$	$3, 6, 6, 12, 6, 12$
17.	18	$0, 1, 2, 4, 5, 8$	$3, 6, 9, 6, 9, 9$
18.	18	$0, 1, 2, 4, 5, 7$	$3, 9, 6, 9, 6, 9$
19.	16	$0, 1, 2, 3, 7, 15$	$4, 4, 4, 8, 16, 16$
20.	12	$0, 1, 2, 3, 7, 11$	$4, 4, 4, 12, 12, 12$
21.	8	$0, 1, 2, 3, 6, 7$	$4, 4, 8, 8, 8, 8$
22.	12	$0, 1, 2, 3, 5, 11$	$4, 6, 4, 6, 12, 12$
23.	8	$0, 1, 2, 3, 5, 7$	$4, 8, 4, 8, 8, 8$
24.	24	$0, 1, 2, 3, 5, 6$	$4, 6, 8, 6, 6, 8$
25.	10	$0, 1, 2, 3, 4, 9$	$5, 5, 5, 5, 10, 10$
26.	6	$0, 1, 2, 3, 4, 5$	$6, 6, 6, 6, 6, 6$

Tabulka 2: Výstup pro $\mathcal{A} = \{0, 1, 2, 3, 4, 5\}$