

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Separating Words Problem

Csányi Dávid

Eötvös Loránd Tudományegyetem

Témavezető: Pálvölgyi Dömötör

2022. december

Egy egyszerű számítási modellel milyen nehéz megoldani a legkönnyebb feladatot, két szó megkülönböztetését?

- $\text{sep}_I(a_1, \dots, a_I)$ az a legkisebb k szám, amire létezik egy k állapotú véges automata, ami az a_1, a_2, \dots, a_I szavak esetén különböző végállapotokba kerül
- $S_I(n) = \max\{\text{sep}_I(a_1, \dots, a_I) : a_1, \dots, a_I \text{ a } \Sigma \text{ abc feletti maximum } n \text{ hosszú páronként különböző szavak}\}$

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Felső korlátok:

- 1986 Goralcik, Koubek: kérdés megfogalmazása, $S_2(n) = o(n)$
- 1989 Robson: $S_2(n) = \mathcal{O}(n^{\frac{2}{5}}(\log n)^{\frac{3}{5}})$
- 2020 Chase: $S_2(n) = \mathcal{O}(n^{\frac{1}{3}}(\log n)^7)$

A legjobb ismert alsó korlát $S_2(n) = \Omega(\log(n))$

Különböző hosszú szavak

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Lemma

[2] Ha $0 \leq i, j \leq n$, $n \geq 2$ és $i \neq j$, akkor létezik egy p prím, amelyre $p \leq 4,4 \cdot \log(n)$ és $i \not\equiv j \pmod{p}$.

Állítás

[1] Ha $|a|, |b| \leq n$ és $|a| \neq |b|$, akkor $\text{sep}_2(a, b) = \mathcal{O}(\log(n))$.

Különböző hosszú szavak

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Állítás

Ha $|a|, |b|, |c| \leq n$ és páronként különbözőek, akkor
 $\text{sep}_3(a, b, c) = \mathcal{O}(\log(n))$.

Állítás

$\text{sep}_3(a, b, c) \leq \min H \cdot \max H$, ahol
 $H = \{\text{sep}_2(a, b), \text{sep}_2(b, c), \text{sep}_2(c, a)\}$

Állítás

Ha a, b és c közül valamelyik szó hossza nem n , akkor
 $\text{sep}_3(a, b, c) \leq S_2(n) \cdot \mathcal{O}(\log(n))$.

Az abc mérete

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Állítás

$$[1] S_2^\kappa(n) = S_2^2(n), \text{ ha } \kappa \geq 2.$$

Állítás

$$S_3^\kappa(n) = S_3^2(n), \text{ ha } \kappa \geq 2.$$

Ötlet: egy $\Phi : \Sigma \mapsto \{0, 1\}$ függvény segítségével egy Σ abc feletti szó párt binárisá alakítunk úgy hogy különbözőek maradjanak.

Speciális esetek

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Bizonyos esetekben könnyen megkülönböztethető az a, b szó pár([1]):

- Eltérés az i -edik helyen $\rightarrow sep_2(a, b) \leq i + 2$
- Eltérés a hátulról i -edik helyen $\rightarrow sep_2(a, b) \leq i + 1$
- Egy d hosszúságú minta előfordulásainak száma eltérő a két szóban $\rightarrow sep_2(a, b) = \mathcal{O}(d \log(n))$
- A Hamming-távolság kisebb, mint d
 $\rightarrow sep_2(a, b) = \mathcal{O}(d \log(n))$

Eltérések a szavak elején

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Állítás

[1] Ha az a, b szópárra igaz, hogy különböznek az i -edik helyen, akkor $sep_2(a, b) \leq i + 2$.

Állítás

Ha az $a_1, a_2, \dots, a_l \in \{0, 2, \dots, \kappa - 1\}^n$ szósorozatra igaz, hogy bármely két szó különbözik az első i hely valamelyikén, akkor $sep_l^\kappa(a_1, \dots, a_l) \leq 1 + \lfloor \frac{l}{2} \rfloor (i - 1) + l = 1 + \lfloor \frac{l}{2} \rfloor + \lfloor \frac{l}{2} \rfloor i$.

Átlagos eset

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Állítás

[1] Ha az a, b szópárt egyenletes eloszlás szerint véletlen választjuk a $\kappa \geq 2$ méretű abc feletti, n hosszú, különböző szópa-
rok halmazából, akkor $\text{sep}_2^\kappa(a, b)$ várható értéke konstans.

Állítás

Ha az a_1, a_2, \dots, a_l szószorozat egyenletes eloszlás szerint véletlen választjuk a $\kappa \geq 2$ méretű abc feletti, n hosszú, páronként különböző szavakból álló l hosszú sorozatok halmazából, akkor $\text{sep}_l^\kappa(a_1, \dots, a_l)$ várható értéke konstans.

Tesztek

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

Az $S_2(n)$ és $S_3(n)$ kis n értékek esetén való kiszámítására készítettem egy programot.

Az összes n hosszú szópár vagy szóhármass esetén meghatározza a sep_2 vagy sep_3 értéket a lehetséges nem izomorf automaták kipróbálásával.

Az alábbi táblázatban láthatók a kiszámolt értékek:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\lfloor \log_2(n) \rfloor$	0	1	2	2	3	3	3	3	4	4	4	4	4	4	4
$S_2(n)$	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4
$S_3(n)$	-	3	3	3	4	4	4	5	5						

Hivatkozások

Csányi Dávid

Bevezetés

Történet

Különböző
hosszú szavak

Az abc mérete

Speciális
esetek

Eltérések a
szavak elején

Átlagos eset

Tesztek

Hivatkozások

- [1] Erik D. Demaine, Sarah Eisenstat, Jeffrey Shallit és David A. Wilson. „Remarks on Separating Words”. (2011). URL: <https://arxiv.org/abs/1103.4513v1>.
- [2] J. Shallit és Y. Breitbart. „Automaticity I: Properties of a measure of descriptive complexity.”. *J. Comput. System Sci.* 53 (1996), 10–25. old.