

Sveučilište u Zagrebu  
Fakultet elektrotehnike i računarstva



# Izgradnja sufiksnog polja korištenjem Kärkkäinen – Sandersovog algoritma

PROJEKT IZ PREDMETA BIOINFORMATIKA

Studenti:

0036449075, Petra Bevandić  
0036449278, Miranda Kreković  
0036449374, Domagoj Šalković  
0036452457, Petra Vučković  
0112019329, Vedran Vukotić

Predavači:

dr.sc. Mirjana Domazet-Lošo  
doc.dr.sc. Mile Šikić

# Sadržaj

- ▶ uvod
- ▶ algoritam
- ▶ implementacija
- ▶ testiranje
- ▶ rezultati

# Uvod

- ▶ sufiksna stabla
  - ▶ svi mogući sufiksi
  - ▶ indeksiranje tekstova, traženje uzoraka, usporedba podnizova
  - ▶ memorijski zahtjevi
- ▶ sufiksna polja

8	→	ACC
5	→	ACCACC
1	→	ACCCACCACC
10	→	C
7	→	CACC
4	→	CACCACC
9	→	CC
6	→	CCACC
3	→	CCACCACC
2	→	CCCACCACC
0	→	GACCCACCACC

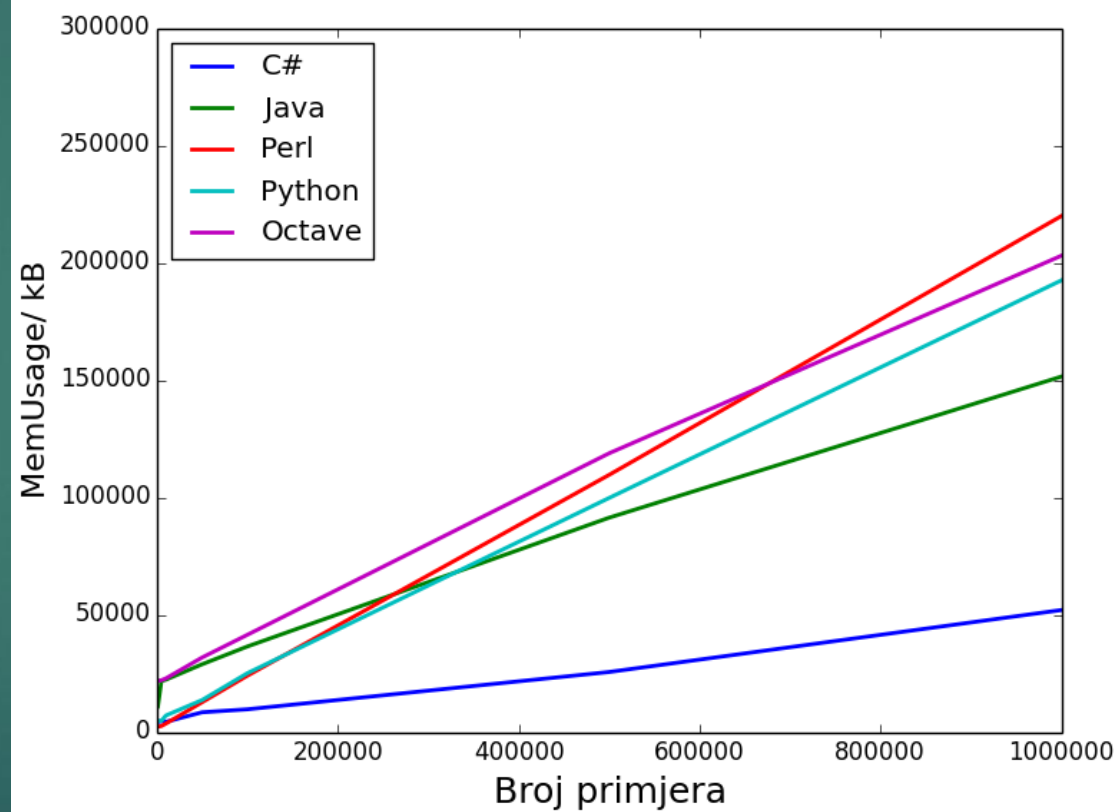
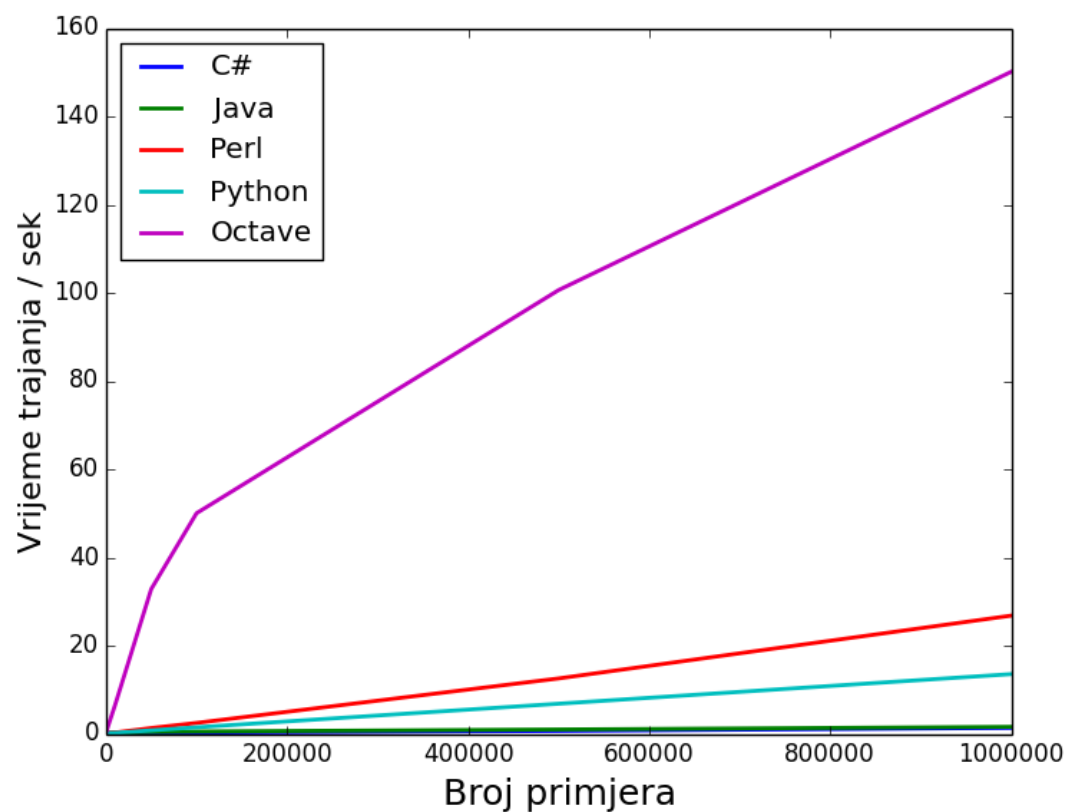
# Algoritam

- ▶ linearno vrijeme izvođenja
- ▶ radix sort

# Implementacija

- ▶ Jezici:
  - ▶ C#
  - ▶ Java
  - ▶ Octave
  - ▶ Perl
  - ▶ Python

# Rezultati



Pitanja...

# Primjer

G	A	C	C	C	A	C	C	A	C	C
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
3	1	2	2	2	1	2	2	1	2	2

Tekst:	3	1	2	2	2	1	2	2	1	2	2	0	0	0
Indeks:	0	1	2	3	4	5	6	7	8	9	10	11	12	13

Grupa 1				Grupa 2		
1 2 2	2 1 2	2 1 2	2 0 0	2 2 2	1 2 2	1 2 2
1	4	7	10	2	5	8

Grupa 1 i 2						
2 0 0	1 2 2	2 1 2	2 1 2	2 2 2	1 2 2	1 2 2
10	1	4	7	2	5	8

Grupa 1 i 2						
2 0 0	2 1 2	2 1 2	1 2 2	2 2 2	1 2 2	1 2 2
10	4	7	1	2	5	8

Grupa 1 i 2						
1 2 2	2 0 0	2 1 2	2 1 2	1 2 2	1 2 2	2 2 2
1	10	4	7	5	8	2

Time:	Grupa 1				Grupa 2		
	1 2 2	2 1 2	2 1 2	2 0 0	2 2 2	1 2 2	1 2 2
	1	3	3	2	4	1	1



# Primjer

Tekst:	1	3	3	2	4	1	1	0	0	0
Indeks:	0	1	2	3	4	5	6	7	8	9

Grupa 1 i 2				
0 0 0	4 1 1	3 3 2	1 1 0	3 2 4
7	4	1	5	2
0 0 0	4 1 1	3 3 2	1 1 0	3 2 4
7	4	1	5	2
0 0 0	3 3 2	4 1 1	1 1 0	3 2 4
7	1	4	5	2

	Grupa 1			Grupa 2	
Indeks:	3 3 2	4 1 1	0 0 0	3 2 4	1 1 0
Ime:	4	5	1	3	2

Indeks:	0 0 0	1 1 0	3 2 4	3 3 2	4 1 1
SA:	2	4	3	0	1

	Grupa 0			Grupa 12				
Indeks:	1 0 0	1 3 3	2 4 1	0 0 0	1 1 0	3 2 4	3 3 2	4 1 1
Redni br:	6	0	3	7	5	2	1	4
SA0/SA12:	6	0	3	2	4	3	0	1

SA	6	5	0	3	2	1	4
----	---	---	---	---	---	---	---

# Primjer

	Grupa 1				Grupa 2		
	1 2 2	2 1 2	2 1 2	2 0 0	2 2 2	1 2 2	1 2 2
Indeks:	1	4	7	10	2	5	8
Ime:	3	6	5	4	7	2	1

	Grupa 0				Grupa 12							
	2 2 0	2 2 1	2 2 1	3 1 2	1 2 2	1 2 2	1 2 2	2 0 0	2 1 2	2 1 2	2 2 2	
Indeks:	9	6	3	0	8	5	1	10	7	4	2	
SA0/SA12	1	2	3	4	1	2	3	4	5	6	7	
	9	6	3	0	6	5	0	3	2	1	4	

SA0/SA12	8	5	1	10	7	4	9	6	3	2	0
----------	---	---	---	----	---	---	---	---	---	---	---