Algoritmi in podatkovne strukture 1

Visokošolski strokovni študij Računalništvo in informatika

Napredni algoritmi urejanja zaporedja



Jurij Mihelič, UniLj, FRI

Napredna urejanja

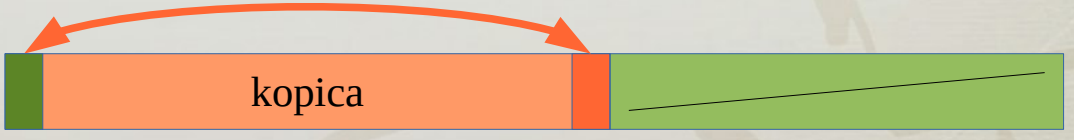
- Napredna urejanja
  - urejanje s **kopico** (*heap* sort)
  - urejanje z **zlivanjem** (*merge* sort)
  - **hitro** urejanje (*quicksort*)
  - …

Urejanje s kopico

- Ideja algoritma (*heapsort*)
  - nadgradimo urejanje z izbiranjem
  - za iskanje pravega elementa uporabimo kopico
  - 1. poskus:
    - prvi del polja je urejeni seznam
    - drugi del polja je kopica

Urejanje s kopico

- Ideja algoritma (*heapsort*)
  - nadgradimo urejanje z izbiranjem
  - za iskanje največjega elementa uporabimo kopico
  - kopico zgradimo v prvem delu tabele
  - drugi del tabele je urejeni del
  - ponavljamo
    - zamenjaj koren kopice z zadnjim elementom kopice
    - ugreznemo koren
- Sled



Urejanje s kopico

- Psevdokoda
- Zahtevnost: $O(n \lg n)$



Urejanje s kopico

```
fun heapSort(a) is
    ;; zgradi kopico
    for i = n / 2 - 1 to 0
        siftDown(a, i)

    ;; urejaj
    while last => 1 do
        swap(a, 0, last)
        last -= 1
        siftDown(0)
```

John von Neumann, 1903-1957



Urejanje z zlivanjem

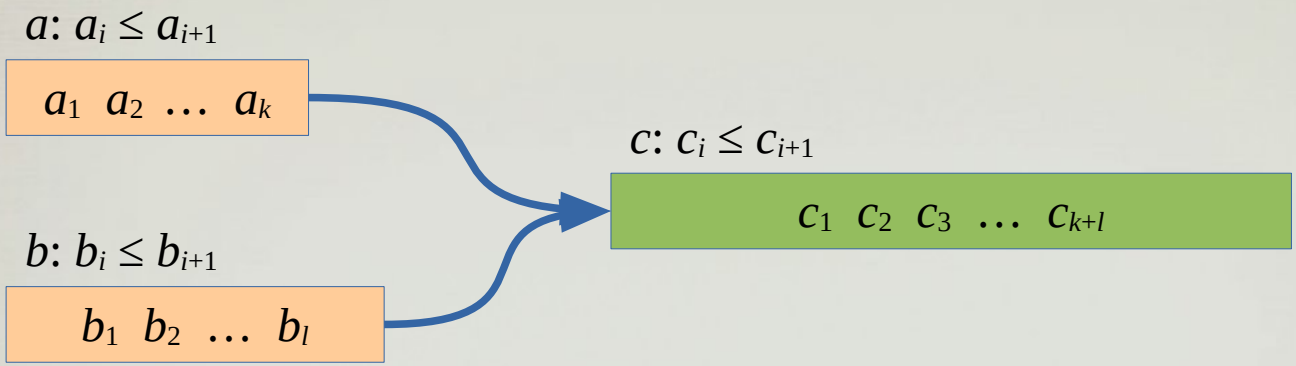
- Ideja algoritma – deli & vladaj
  - tabelo razdelimo na dve polovici
  - rekurzivno uredimo obe podtabeli
  - zlijemo obe urejeni podtabeli





Urejanje z zlivanjem

- Zlivanje urejenih podtabel



- Ideja algoritma
  - hkratni zaporedni sprehod po zaporedjih
- Zahtevnost zlivanja
  - $\Theta(k+l)$

Urejanje z zlivanjem

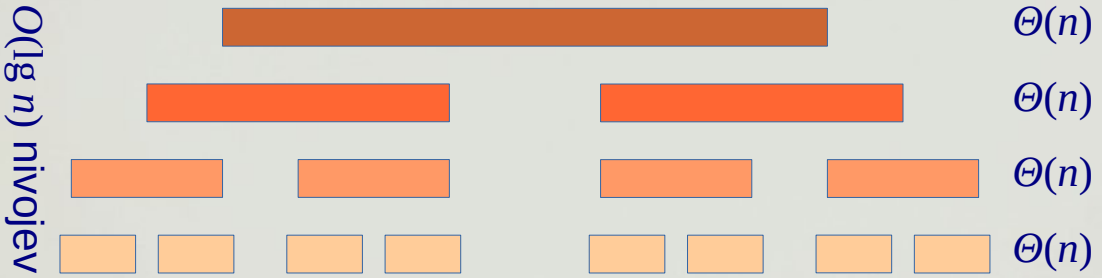
- Psevdokoda

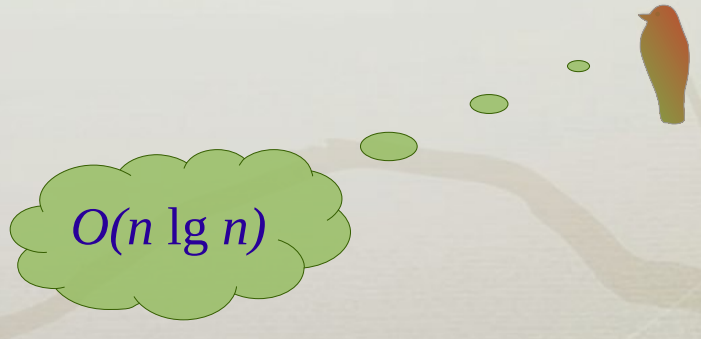
Urejanje z zlivanjem

```
fun mergesort(a) is
    if a.length <= 1 then return a
    middle = (a.length - 1) / 2
    left = mergesort(a[0 ... middle])
    right = mergesort(a[middle+1 ... a.length-1])
    return merge(left, right)
end
```

Urejanje z zlivanjem

- Zahtevnost algoritma
  - Kako globoka je lahko največ rekurzija?
  - Koliko dela je **v celoti** na vsakem nivoju rekurzije?





Hitro urejanje

C. A. R. Hoare, 1934

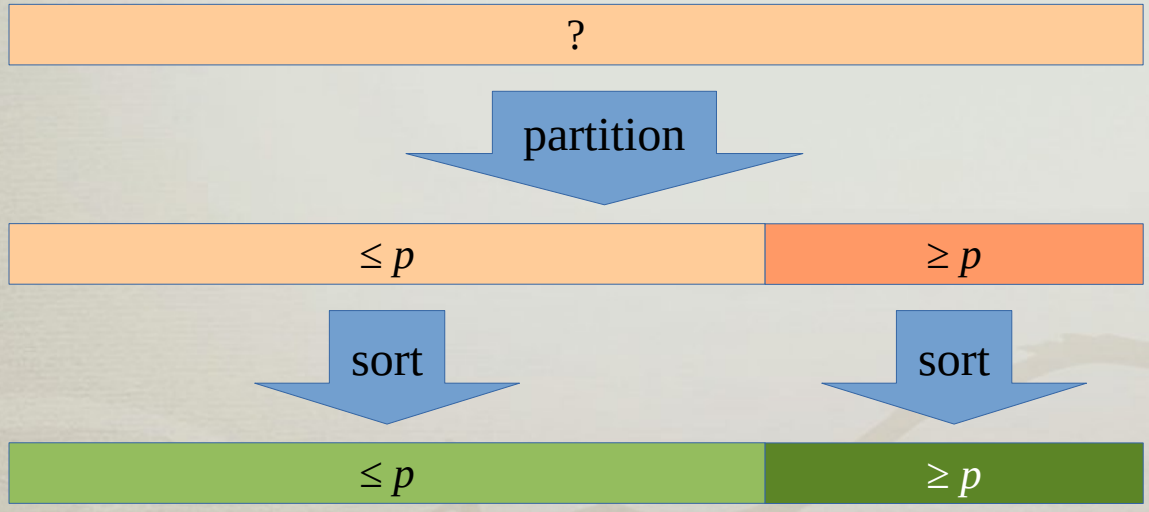


Quicksort, 1960

- Quicksort
  - eden izmed najpogosteje uporabljanih algoritmov za urejanje
  - splošen, uporablja primerjave
  - algoritem deluje neposredno v tabeli
  - potrebuje malo dodatnega prostora
  - dobro deluje za različne vrste podatkov
  - v povprečju zelo hiter
  - previdno pri implementaciji

Hitro urejanje

- Ideja algoritma – deli & vladaj
  - tabelo porazdelimo na dva dela:
    - s pomočjo poljubnega elementa p (pivot)
    - levi del vsebuje elemente $\leq p$ in desni elemente $\geq p$
  - rekurzivno uredimo obe podtabeli



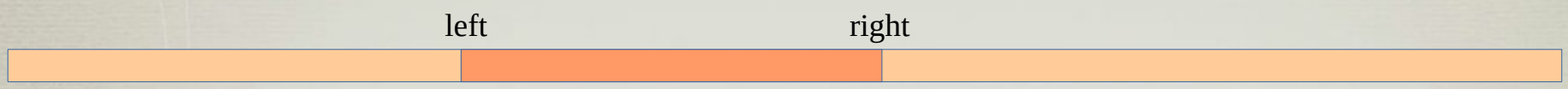


Hitro urejanje

- Porazdeljevanje
  - veliko načinov
    - izbira pivota
    - z dodatnim poljem
    - križanje kazalcev
    - enozančno
    - uporaba čuvajev
    - pogoji v zankah
    - ipd.

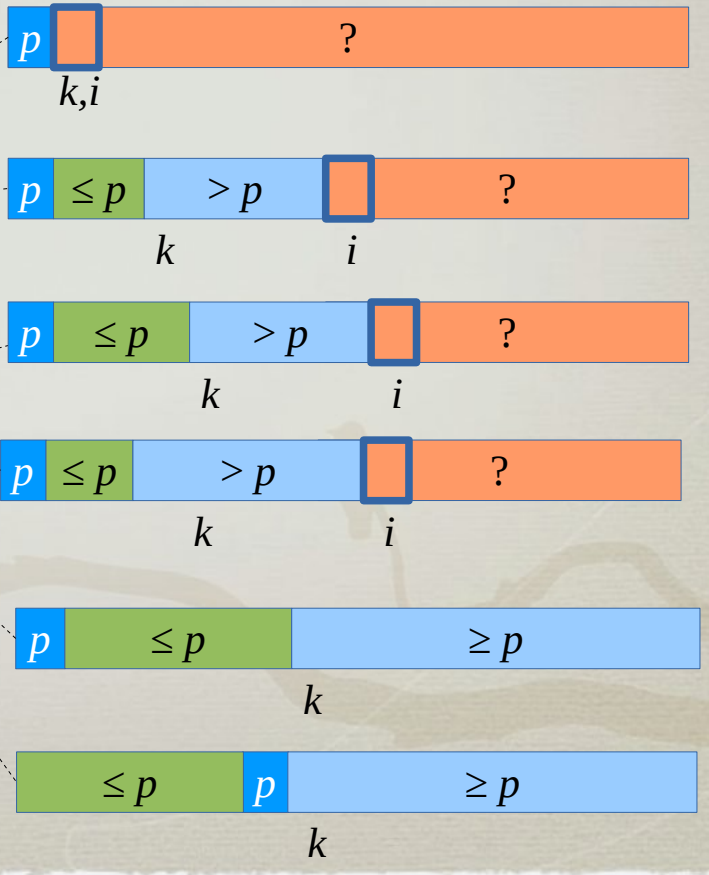
Hitro urejanje

- Porazdeljevanje



Porazdeljevanje

```
p = a[left]
k = left + 1
for i = left + 1 to right do
    if a[i] <= p then
        swap(a, k, i);
        k += 1
    endif
endfor
swap(a, left, k-1)
return k - 1
```



Hitro urejanje

Hitro urejanje

```
fun partition(a, left, right) is
    p = a[left]
    k = left + 1
    for i = left + 1 to right do
        if a[i] <= p then
            swap(a, k, i);
            k += 1
        endif
    endfor
    swap(a, left, k-1)
    return k - 1
end

fun quicksort(a, left, right) is
    if left >= right then return
    r = partition(a, left, right)
    quicksort(a, left, r - 1)
    quicksort(a, r + 1, right)
end
```

Hitro urejanje

- Sled
- Zahtevnost
  - best: $O(n \lg n)$
  - worst: $O(n^2)$
  - average: $O(n \lg n)$
- Prostorska zahtevnost
  - $O(n)$
  - skrbna implementacija: $O(\lg n)$
    - na sklad damo večji interval
    - repno rekurzijo spremenimo v zanko

Hitro urejanje

- Realni podatki
  - pogosto že (delno) urejena zaporedja
  - zahtevnost lahko blizu najslabše
- Izbira pivota
  - levi, desni, srednji element
  - mediana treh, petih, itd.
  - randomizacija – naključni pivot
    - `swap(a, left, random)`
    - `pivot = a[random]`
      - prekoračitev obsega, pazljivo pri izvedbi
    - pričakovana zahtevnost: $O(n \lg n)$



Inženiring algoritmov

- Stabilnost
  - vstavljanje (IS), zamenjave (BS), zlivanje (MS)
- Velikost zaporedja
  - majhno: vstavljanje (IS)
  - veliko: veliki trije (QS, MS, HS)
- Rekurzija
  - nikoli do konca (MS, QS)
  - ustavimo z vstavljanjem (IS)



Inženiring algoritmov

- State of the art
  - 2002, TimSort: MS+IS, python/java
  - 2009, Jaroslavski dvo-pivotno hitro urejanje
    - Java za primitivne tipe, ustavljanje QS z IS pri n=47
  - 2014, 5-pivotno hitro urejanje, predpomnilnik
    - Kushagra, Lopez-Ortiz, Munro, Qiao
- Zunanje urejanje
  - kadar tabela ne gre v pomnilnik
  - zlivanje: navadno/naravno, …, polifazno, kaskadno

Povzetek

Vrsta urejanja	Zahtevnost	Razno
Navadno vstavljanje	$O(n^2)$, best: $O(n)$	stabilno
Navadno izbiranje	$\Theta(n^2)$	
Navadne zamenjave	$\Theta(n^2)$	stabilno
Urejanje s kopico	$\Theta(n \log n)$	
Urejanje z zlivanjem	$\Theta(n \log n)$	stabilno, ni *in-place*, dodatni prostor
Hitro urejanje	$O(n^2)$, avg: $\Theta(n \log n)$	randomizacija, dodatni prostor
Urejanje s štetjem	$O(n + m)$	stabilno, končna množica
Korensko urejanje	$O(d(n + m))$	stabilno, končna množica
Urejanje s koši	$O(n^2)$, avg: $\Theta(n)$	stabilno?, enakomerno