Proiectarea algoritmilor: complexitatea medie

Ștefan Ciobâcă, Dorel Lucanu

Faculty of Computer Science
Alexandru Ioan Cuza University, Iași, Romania
dlucanu@info.uaic.ro

PA 2015/2016

1. Timpul mediu: algoritmi determiniști
   - Quicksort determinist
2. Timpul mediu: algoritmi probabiliști
   - Agoritmi nedeterminiști în general
   - Agoritmi nedeterminiști pentru probleme de decizie
   - Algoritmi probabiliști
   - Quicksort probabilist
3. k-mediana

Plan

1. Timpul mediu: algoritmi deterministi
   - Quicksort determinist
2. Timpul mediu: algoritmi probabilisti
   - Agoritmi nedeterministi în general
   - Agoritmi nedeterministi pentru probleme de decizie
   - Algoritmi probabilisti
   - Quicksort probabilist
3. *k*-mediana

Motivație

- A rezolvă P, $x \in P$
- timpul în cazul cel mai nefavorabil:
  $T_A(n) = \sup\{time(A,p) \mid x \in P \wedge size(x) = n\}$
- $T_A(n)$ este irelevant dacă numărul instanțelor x cu $size(x) = n \wedge time(A,x) = T_A(n)$ (sau $time(A,x) = T_A(n) - \varepsilon$) este foarte mic

Definiție

- $time(A, x)$ ca variabilă aleatorie:
  - o experiență = execuția algoritmului pentru o instanță x,
  - valoarea experienței = durata execuției algoritmului
- legea de repartiție a acestei variabile aleatorie
- timpul de execuție mediu = media acestei variabile aleatoare

$$T_A^{med}(n) = M(\{time(A, x) \mid x \in P \wedge size(x) = n\})$$

- Caz particular: $time(A, x) = \{t_0, t_1, \dots\}$,
  $Pr(x \mid time(A, x) = t_i) = p_i$

$$T_A^{med}(n) = \sum_i t_i \cdot p_i$$

$Pr(x \mid time(A, x) = t_i)$: probabilitatea ca timpul de execuție al lui A pentru intrarea x să fie t_i

Exemplu

Problema FIRST OCCURRENCE

Input: $n, a = (a_0, \ldots, a_{n-1}), z$, toate numere întregi.

Output: $poz = \begin{cases} \min\{i \mid a_i = z\} & \text{dacă } \{i \mid a_i = z\} \neq \emptyset, \\ -1 & \text{altfel.} \end{cases}$

Algoritm pentru FIRST OCCURRENCE

Algoritmul FOAlg descris de următorul program rezolvă FIRST OCCURRENCE:

```
//@input: un tablou a cu n elemente, z
//@output:  pozitia primului element din a egal cu z,
//          -1 daca nu exista un astfel de element
i =  0;
while (a[i] != z) && (i < n-1) {
  i = i+1;
if (a[i] == z)  poz = i;
else poz = -1;
```

Timpul mediu pentru FOAlg 1/2

dimensiune instanță: $n = \texttt{a.size()}$

operații măsurate: atribuiri și comparații

$time(\text{FOAlg}, \{x \mid size(x) = n\}) = \{3i + 2 \mid 1 \le i \le n\}$

Presupuneri:

- probabibilitatea ca $z \in \{a_0, \ldots, a_{n-1}\}$ este q și
- probabilitatea ca z să apară prima dată pe poziția $i - 1$ este $\frac{q}{n}$ (indicii i candidează cu aceeași probabilitate pentru prima apariție a lui z).

Timpul mediu pentru FOAlg 2/2

$$Pr(z \notin \{a_0, \ldots, a_{n-1}\}) = 1 - q$$

$$Pr(time(\mathrm{FOAlg}, p) = 3i + 2) = \frac{q}{n} = p_i,\ 1 \leq i < n$$

$$Pr(time(\mathrm{FOAlg}, p) = 3n + 2) = \frac{q}{n} + (1 - q) = p_n$$

Timpul mediu de execuție este:

$$\begin{aligned} T^{med}_{\mathrm{FOAlg}}(n) &= \sum_{i=1}^{n} p_i x_i \\ &= \sum_{i=1}^{n-1} \frac{q}{n} \cdot (3i + 2) + (\frac{q}{n} + (1 - q)) \cdot (3n + 2) \\ &= 3n - \frac{3nq}{2} + \frac{3q}{2} + 2 \end{aligned}$$

Plan

1. Timpul mediu: algoritmi deterministi
   - Quicksort determinist
2. Timpul mediu: algoritmi probabilisti
   - Agoritmi nedeterministi în general
   - Agoritmi nedeterministi pentru probleme de decizie
   - Algoritmi probabilisti
   - Quicksort probabilist
3. *k*-mediana

Quicksort: descriere

Este proiectat pe paradigma divide-et-impera.

Algoritmul Quicksort

Input: $S = \{a_0, \ldots, a_{n-1}\}$

Output: o secvență cu elementele a_i în ordine crescătoare

1. se alege $x \in S$
2. calculează
   $S_< = \{a_i \mid a_i < x\}$ $S_= = \{a_i \mid a_i = x\}$ $S_> = \{a_i \mid a_i > x\}$
3. sortează recursiv $S_<$ și $S_>$ producând $Seq_<$ și $Seq_>$, respectiv
4. întoarce secvența $Seq_<$, $S_=$, $Seq_>$

Quicksort: partiționarea

Presupunem că S este memorată într-un tablou a.
Următoarea soluție utilizează un tablou suplimentar:

```
partition(out a, p, q, out k) {
  l = 0;
  for (i = p; i <= q; ++i) {
     if (a[i] <= a[p]) {
      b[l] = a[i]; l = l + 1;
    }
  }
  k = l - 1;
  for (i = p; i <= q; ++i) {
     if (a[i] > a[p]) {
      b[l] = a[i]; l = l + 1;
    }
  }
  a[p..q] = b;   // pseudocode
}
```

Quicksort: partiționarea, eliminarea tabloului auxiliar, 1/5

Se determină prin interschimbări a unui indice k cu proprietățile:

- $p \leq k \leq q$ și $a[k] = x$;
- $\forall i : p \leq i \leq k \implies a[i] \leq a[k]$;
- $\forall j : k < j \leq q \implies a[k] \leq a[j]$;

- inițial: $x = a[p]$, $i = p + 1$ și $j = q$
- proprietățile menținute invariante:

$$\forall i' : p \leq i' < i \implies a[i'] \leq x \tag{1}$$

și

$$\forall j' : j < j' \leq q \implies a[j'] \geq x \tag{2}$$

Quicksort: partiționarea, eliminarea tabloului auxiliar, 2/5

Presupunem că la momentul curent sunt interogate elementele a[i] și a[j] cu $i < j$. Distingem următoarele cazuri:

1. $a[i] \leq x$. Transformarea `i = i+1` păstrează 1.
2. $a[j] \geq x$. Transformarea `j = j-1` păstrează 2.
3. $a[i] > x > a[j]$. Dacă se realizează interschimbarea a[i] $\leftrightarrow$ a[j] și se face `i = i+1` și `j = j-1`, atunci ambele predicate (1) și (2) sunt păstrate.

Quicksort: partiționarea, eliminarea tabloului auxiliar, 3/5

Operațiile de mai sus sunt repetate până când *i* devine mai mare decât *j*:

```
while (i ≤ j) {
  if (a[i] ≤ x) i = i+1;
  else if (a[j] ≥ x) j = j-1;
  else if ((a[i] > x) && (x > a[j])) {
    swap(a, i, j);
    i = i+1;
    j = j-1;
  }
}
```

Quicksort: partiționarea, eliminarea tabloului auxiliar, 4/5

Analiza terminării lui `while`:

- $i = j + 1$:
- din 1, 2 avem $a[i-1] \leq x$ și $a[i] = a[j+1] \geq x$
- deci interschimbând a[p] cu a[i − 1] obținem partiționarea dorită a tabloului $\implies k = i - 1$

```
k = i-1;  a[p] = a[k];  a[k] = x;
```

- analiza cazurilor limită:
  $i = p + 1$ – relațiile de mai sus au sens
  $j = q \implies k = q$, i.e. $a[p..k] = a[p..q]$ ce ar putea conduce la recursie infinită; în acest caz k trebuie decrementat

Quicksort: algoritmul de partiționare, eliminarea tabloului auxiliar, 5/5

```
@input:   a = (a[p], ..., a[q])
@output: k, a  cu proprietatea
  ∀ i : p ≤ i ≤ k ⟹ a[i] ≤ a[k] şi ∀ j : k < j ≤ q ⟹ a[k] ≤ a[j]
partition(out a, p, q, out k) {
  x = a[p] ;
  i = p + 1;   j = q;
  while (i ≤ j) {
    if (a[i] ≤ x) i = i+1;
    else if (a[j] ≥ x) j = j-1;
    else if ((a[i] > x) && (x > a[j])) {
      swap(a, i, j);
      i = i+1;
      j = j-1;
    }
  }
  k = i-1;  a[p] = a[k];  a[k] = x;
//  if (j == q) --k;
}
```

Quicksort: algoritm

@input: $a = (a[p], \ldots, a[q])$
@output: elementele secvenţei a în ordine crescătoare

```
qsort(out a, p, q) {
  if (p < q) {
    partition(a, p, q, k)
    qsort(a, p, k-1)
    qsort(a, k+1, q)
  }
}
```

Quicksort: timpul în cazul cel mai nefavorabil

- dimensiune instanță: $\mathtt{n} = \mathtt{a.size}()$
- operații măsurate: comparații care implică elementele tabloului
- cazul cel mai nefavorabil: tabloul deja ordonat
- numărul de comparații pentru acest caz:
  $(n-1) + (n-2) + \cdots + 1 = O(n^2)$

Quicksort: timpul mediu

- lungimea secvenței: $q + 1 - p = n$
- probabilitatea ca x sa fie al k-lea element: $\frac{1}{n}$
- dimensiuni subprobleme: $k - p = i - 1$ și $q - k = n - i$
- numărul mediu de comparații:

$$T^{med}(n) = \begin{cases} (n-1) + \frac{1}{n}\sum_{i=1}^{n}(T^{med}(i-1) + T^{med}(n-i)) & \text{, dacă } n \geq 1 \\ 1 & \text{, dacă } n = 0 \end{cases}$$

Theorem

Complexitatea medie a algoritmului QuickSort este $O(n \log_2 n)$.

Plan

Plan

1. Timpul mediu: algoritmi determiniști
   - Quicksort determinist
2. Timpul mediu: algoritmi probabiliști
   - Agoritmi nedeterminiști în general
   - Agoritmi nedeterminiști pentru probleme de decizie
   - Algoritmi probabiliști
   - Quicksort probabilist
3. *k*-mediana

Extensia limbajului

`choose` x `in` S`;` – întoarce un element din S ales arbitrar

`choose` x `in` S `s.t.` B`;` – întoarce un element din S care satisface condiția B ales arbitrar

`failure;` – semnalează terminarea fără succes (e.g., o instrucțiune `choose` nu s-a putut executa)

Demo cu noile instructiuni

```
choose x1 in { 1 ..  5 };
```

```
$ ~/k-3.6/bin/krun tests/choose.alk -cINIT=".Map"
<k>
    .K
</k>
<state>
    x1 |-> 3
</state>
$ ~/k-3.6/bin/krun tests/choose.alk -cINIT=".Map"
<k>
    .K
</k>
<state>
    x1 |-> 1
</state>
```

Demo cu noile instructiuni

```
choose x1 in { 1 ..  5 };
```

```
$ ~/k-3.6/bin/krun tests/choose.alk -cINIT=".Map" --search
Search results:
Solution 1:
<k>
    .K
</k>
<state>
    x1 |-> 1
</state>
...
Solution 5:
<k>
    .K
</k>
<state>
    x1 |-> 5
</state>
```

Demo cu noile instructiuni

```
odd(x) {
  return x % 2 == 1;
}
choose x1 in { 1 .. 5 } s.t. odd(x1);
```

```
$ ~/k-3.6/bin/krun tests/choosest.alk -cINIT=".Map"
<k>
    .K
</k>
<state>
    x1 |-> 5
</state>
$ ~/k-3.6/bin/krun tests/choosest.alk -cINIT=".Map"
<k>
    .K
</k>
<state>
    x1 |-> 1
</state>
```

Demo cu noile instructiuni

```
$ ~/k-3.6/bin/krun tests/choosest.alk -cINIT=".Map" --search
Search results:
Solution 1:
<k>
    .K
</k>
<state>
    x1 |-> 1
</state>
Solution 2:
<k>
    .K
</k>
<state>
    x1 |-> 3
</state>
Solution 3:
<k>
    .K
</k>
<state>
    x1 |-> 5
</state>
```

Demo cu noile instructiuni

```
odd(x) {
  return x % 2 == 1;
}

s = emptySet;
for (i = 0; i < 8; i = i+2)
  s.pushBack(i);
choose x in s s.t. odd(x);

$ ~/k-3.6/bin/krun tests/failure.alk -cINIT=".Map"
<k>
    failure ;
</k>
<state>
    i |-> 8
    s |-> { 0, 2, 4, 6 }
    x |-> 6
</state>
```

Problemă rezolvată de un program nedeterminist

- un program nedeterminist are mai multe fire de execuție
- un program nedeterminist rezolvă P dacă $\forall x \in P$ $\exists$ un fir de execuție care se termină și a cărui configurție finală include $P(x)$

Exemplu: problema celor N regine

Input: o tablă de șah $n \times n$.
Output: o așezare a n piese de tip regină pe tablă a.î. nicio regină nu atacă o altă regină.

```
attacked(i, j, b) {
  attack = false;
  for (k = 0;  k < i; ++k)
     if ((b[k] == j) || ((b[k]-j) == (k-i)) || ((b[k]-j) == (i-k)))
       attack = true;
  return(attack);
}

nqueens (n) {
   for (i = 0; i < n; ++i) {
     choose j in { 0 .. n-1 } s.t. ! (attacked(i, j, b));
     b[i] = j;
    }
}
```

Exemplu: problema celor N regine

```
$ ~/k-3.6/bin/krun tests/nqueens.alk -cINIT="n |-> 4"
<k>
    failure ;
</k>
<state>
    b |-> [ 0, 3, 1, -1 ]
    i |-> 3
    j |-> 3
    n |-> 4
</state>
<stack>
    .List
</stack>
```

Exemplu: problema celor N regine

```
$ ~/k-3.6/bin/krun tests/nqueens.alk -cINIT="n |-> 4" --search
Search results:
Solution 1:
<k>
    .K             i.e. success
</k>
<state>
    b |-> [ 1, 3, 0, 2 ]
    n |-> 4
</state>
Solution 2:
<k>
    .K             i.e. success
</k>
<state>
    b |-> [ 2, 0, 3, 1 ]
    n |-> 4
</state>
```

Exemplu: problema celor N regine

```
Solution 3:
<k>
    failure ;
</k>
<state>
    b |-> [ 0, 2, -1, -1 ]
    i |-> 2
    j |-> 3
    n |-> 4
</state>

Solution 4:
<k>
    failure ;
</k>
<state>
    b |-> [ 0, 3, 1, -1 ]
    i |-> 3
    j |-> 3
    n |-> 4
</state>
```

Exemplu: problema celor N regine

```
Solution 5:
<k>
    failure ;
</k>
<state>
    b |-> [ 3, 0, 2, -1 ]
    i |-> 3
    j |-> 3
    n |-> 4
</state>

Solution 6:
<k>
    failure ;
</k>
<state>
    b |-> [ 3, 1, -1, -1 ]
    i |-> 2
    j |-> 3
    n |-> 4
</state>
```

Plan

1. Timpul mediu: algoritmi determiniști
   - Quicksort determinist
2. Timpul mediu: algoritmi probabiliști
   - Agoritmi nedeterminiști în general
   - Agoritmi nedeterminiști pentru probleme de decizie
   - Algoritmi probabiliști
   - Quicksort probabilist
3. *k*-mediana

Definiție

- Activitatea unui algoritm nedeterminist pentru o problemă de decizie:
  - "se ghicește" o anumită structură S
  - verifică dacă S satisface o proprietățile din *output*
- Extensia limbajului:

  `success` – semnalează terminarea verificării (și a a algoritmului) cu succes

Agoritmi nedeterminiști pentru probleme de decizie: exemplu

SAT

Instance: O mulțime finită de variabile și o formulă propozițională F în formă normală conjunctivă. *Question*: Este F adevărată pentru o anume atribuire de variabile? (i.e., este F satisfiabilă?)

```
// guess
for (i = 0; i < n; ++i) {
  choose z in {false, true};
  x[i] = z;
}
// check
if (f(x)) success;
else failure;
```

Exemplu de instanță SAT

```
f(x) {
  return (x[0] || x[1]) &&
         (!x[0] || x[3] || x[2]) &&
         (x[2] || !x[3]) &&
         (!x[1] || !x[2] || x[3]);
}
```

Execuție nedeterministă

```
$ ~/k-3.6/bin/krun tests/sat.alk -cINIT="n |-> 4"
<k>
    failure ;
</k>
<state>
    i |-> 4
    n |-> 4
    x |-> [ false, true, false, true ]
    z |-> true
</state>
<stack>
    .List
</stack>
```

Execuție nedeterministă

```
$ ~/k-3.6/bin/krun tests/sat.alk -cINIT="n |-> 4"
<k>
    failure ;
</k>
<state>
    i |-> 4
    n |-> 4
    x |-> [ false, false, true, true ]
    z |-> true
</state>
<stack>
    .List
</stack>
```

Execuție exhaustivă

```
$ ~/k-3.6/bin/krun tests/sat.alk -cINIT="n |-> 4" --search
Search results:

Solution 1:
<k>
    failure ;
</k>
...
Solution 12:
<k>
    success ;
</k>
<state>
    i |-> 4
    n |-> 4
    x |-> [ false, true, false, false ]
    z |-> false
</state>
<stack>
    .List
</stack>
```

Plan

1. Timpul mediu: algoritmi deterministi
   - Quicksort determinist
2. Timpul mediu: algoritmi probabiliști
   - Agoritmi nedeterminiști în general
   - Agoritmi nedeterminiști pentru probleme de decizie
   - Algoritmi probabiliști
   - Quicksort probabilist
3. k-mediana

Definiții

Exista două puncte de vedere:

- 1. algoritmul probabilist este văzut ca un algoritm nedeterminist pentru care există o distribuție de probabilitate peste alegerile nedetermiste
- 2. algoritmul probabilist este un algoritm care are o intrare suplimentară ce constă într-o secvență de biți aleatorii;

Această diferență poate fi proiectată în complexitate sau ieșire:

- timpul de execuție văzut ca o variabilă aleatorie
- ieșirea văzută ca o variabilă aleatorie

La acest curs consideră doar prima variantă (numiți și algoritmi Las Vegas).

Timpul mediu de execuție al algoritmilor probabiliști 1/2

Notații:

$prob_{A,x}(C)$ = probabilitatea cu care algoritmul A execută calculul C pentru intrarea x

$time(A, C)$ = timpul necesar lui A ca să execute calculul C (un pic diferit față de cazul determinist)

Timpul mediu de execuție al algoritmilor probabiliști 2/2

timpul mediu de execuție a lui A pentru intrarea x este
$exp\text{-}time(A, x) = M[time] = \sum_C prob_{A,x}(C) \cdot time(A, C)$.

$time(A, \_)$ este variabilă aleatorie.

timpul mediu de execuție a lui A în cazul cel mai nefavorabil este
$exp\text{-}time(A, n) = max\{exp\text{-}time(A, x) \mid g(x) = n\}$

Dacă A este subînțeles din context, atunci scriem numai $exp\text{-}time(n)$ ($exp\text{-}time(x)$) în loc de $exp\text{-}time(A, n)$ (resp. $exp\text{-}time(A, x)$).

Plan

1. Timpul mediu: algoritmi deterministi
   - Quicksort determinist
2. Timpul mediu: algoritmi probabiliști
   - Agoritmi nedeterminiști în general
   - Agoritmi nedeterminiști pentru probleme de decizie
   - Algoritmi probabiliști
   - Quicksort probabilist
3. *k*-mediana

"Randomized Quicksort"

- exemplul canonic pentru algoritmii Las Vegas

Algoritmul RQS
Input: $S = \{a_0, \ldots, a_{n-1}\}$
Output: elementele a_i în ordine crescătoare

1. dacă $n = 1$ întoarce a_0, altfel alege aleatoriu $k \in \{0, \ldots, n-1\}$
2. calculează
   $S_< = \{a_i \mid a_i < a_k\}$
   $S_= = \{a_i \mid a_i = a_k\}$
   $S_> = \{a_i \mid a_i > a_k\}$
3. sortează recursiv $S_<$ și $S_>$ producând $Seq_<$ și $Seq_>$, resp
4. întoarce secvența $Seq_<$, $Seq_=$, $Seq_>$

"Randomized Quicksort"

Se modifică doar algoritmul de partiționare:

```
partition(out a, p, q, out k) {
  l = 0;
  k = p + random(q-p);
  for (i = p; i <= q; ++i) {
     if (a[i] <= a[k]) {
      b[l] = a[i]; l = l + 1;
    }
  }
  k = l - 1;
  for (i = p; i <= q; ++i) {
     if (a[i] > a[k]) {
      b[l] = a[i]; l = l + 1;
    }
  }
  a[p..q] = b;  // pseudocode
}
```

Analiza algoritmului RQS

Fie funcția *rank* astfel încât $a_{rank(0)} \leq \ldots \leq a_{rank(n-1)}$.

Definim $X_{ij} = \begin{cases} 1 & a_{rank(i)} \text{ şi } a_{rank(j)} \text{ sunt comparate} \\ 0 & altfel \end{cases}$

X_{ij} numără comparațiile dintre $a_{rank(i)}$ şi $a_{rank(j)}$

X_{ij} este variabilă aleatorie

Numărul mediu de comparații este

$M[\sum_{i=0}^{n-1} \sum_{j>i} X_{ij}] = \sum_{i=0}^{n-1} \sum_{j>i} M[X_{ij}]$

Analiza algoritmului RQS

p_{ij} probabiltatea ca $a_{rank(i)}$ şi $a_{rank(j)}$ să fie comparate într-o execuție

$M[X_{ij}] = p_{ij} \times 1 + (1 - p_{ij}) \times 0 = p_{ij}$

$p_{ij} = \dfrac{2}{j - i + 1}$

$$
\begin{aligned}
\sum_{i=0}^{n-2} \sum_{j>i} p_{ij} &= \sum_{i=0}^{n-2} \sum_{j>i} \frac{2}{j-i+1} \\
&\leq \sum_{i=0}^{n-2} \sum_{k=1}^{n-i-1} \frac{2}{k} \\
&\leq 2 \sum_{i=0}^{n-2} \sum_{k=1}^{n-i-1} \frac{1}{k}
\end{aligned}
$$

Analiza algoritmului RQS

Theorem

Numărul mediu de comparații într-o execuție al algoritmului RQS este cel mult $2nH_n = O(n \log n)$.

Plan

1. Timpul mediu: algoritmi determiniști
   - Quicksort determinist
2. Timpul mediu: algoritmi probabiliști
   - Agoritmi nedeterminiști în general
   - Agoritmi nedeterminiști pentru probleme de decizie
   - Algoritmi probabiliști
   - Quicksort probabilist
3. k-mediana

k-mediana: problema

Definition

Fie S o listă cu n elemente dintr-o mulțime univers total ordonată. k-mediana este cel de-al k-lea element din lista sortată a elementelor din S.

Presupunem S memorată într-un tablou.

Considerăm următoarea problemă:

Input un tablou $(a[i] \mid 0 \leq i < n)$ și un număr $k \in \{0, 1, \ldots, n-1\}$,

Output k-mediana

k-mediana: descriere algoritm

Mediana poate fi obținută cu algoritmul de partiționare de la `quickSort`.

Condiția pe care trebuie să o satisfacă la ieșire tabloul a este formulată de:

$$(\forall i)(i < k \implies a[i] \leq a[k]) \wedge (i > k \implies a[i] \geq a[k])$$

Fie j poziția calculată de algoritmul de partiționare, i.e:

$$(\forall i)(i < j \implies a[i] \leq a[j]) \wedge (i > j \implies a[i] \geq a[j])$$

1. $j = k \implies$ problema este rezolvată
2. $j < k \implies$ caută k în `a[j+1..n]`
3. $j > k \implies$ caută k în `a[1..j-1]`

k-mediana: algoritmul recursiv

Aceasta conduce la următoarea formulare recursivă a algoritmului de selectare:

```
@input: un tablou a cu n elemente, 0 ≤ k < n
@output: k-mediana
qselect(out a, p, q, k) {
  partition(a, p, q, j);
  if (j == k) return a[k];
  if (j < k) qselect(a, j + 1, q, k);
  else qselect(a, p, j - 1, k);
}
```

k-mediana: algoritmul nerecursiv

Descrierea recursivă nu este avantajoasă deoarece produce un consum de memorie suplimentară (stiva apelurilor recursive) ce poate fi eliminat prin derecursivare:

```
@input: un tablou a cu n elemente, 0 ≤ k < n
@output: k-mediana
qselect(out a, n, k) {
  p = 0;  1 = n-1;
  repeat
    partition(a, p, q, j);
    if (j < k) p = k1 + 1;
    if (k < j) q = k1 - 1;
  until (j == k);
  return a[k];
}
```

qselect: analiza 1/3

Proprietate: Fie dată o bară de lungime 1, care se taie arbitrar în două. Lungimea medie a bucății mai lungi este $\frac{3}{4}$.

Concluzie: daca se împarte aleatoriu în două un tablou de lungime n, lungimea medie celui mare subtablou este $\frac{3}{4}n$.

qselect: analiza 2/3

exp-time(n, k) - timpul mediu pentru a găsi k-mediana

exp-time$(n) = \max_k$ *exp-time*(n, k)

$$
\begin{aligned}
\textit{exp-time}(n) &\leq (n-1) + \frac{2}{n}\sum_{i=\frac{n}{2}}^{n-1} \textit{exp-time}(i) \\
&= (n-1) + \mathrm{avg}[\textit{exp-time}(\frac{n}{2}), \ldots, \textit{exp-time}(n-1)]
\end{aligned}
$$

(un raționament similar ca la qsort)

qselect: analiza 3/3

Afirmăm că $\mathit{exp\text{-}time}(n) \le 4n$. Demonstrăm prin inducție.

Baza: $n = 1$

Pasul inductiv:

Ipoteza inductivă: $\mathit{exp\text{-}time}(i) \le 4i$, $i = n/2, \ldots, n-1$

Avem

$$
\begin{aligned}
\mathit{exp\text{-}time}(n) &\le (n-1) + \mathrm{avg}[\mathit{exp\text{-}time}(\frac{n}{2}), \ldots, \mathit{exp\text{-}time}(n-1)] \\
&\le (n-1) + \mathrm{avg}[4\frac{n}{2}, \ldots, 4(n-1)] \\
&\le (n-1) + 4\frac{3}{4}n \\
&< 4n
\end{aligned}
$$

Un algoritm determinist liniar

1. grupează tabloul în $\frac{n}{5}$ grupe de 5 elemente și calculează mediana fiecărei grupe;
2. calculează recursiv mediana medianelor p
3. utilizează p ca pivot și separă elementele din tablou
4. apelează recursiv pentru subtabloul potrivit (în care se află k-mediana)

Un algoritm determinist liniar: analiza

Notații: $T(n, k)$ timpul pentru cazul cel mai nefavorabil pentru k-mediana,
$T_n = \max_k T(n, k)$
Pasul 1: $O(n)$
Pasul 2: $T(n/5)$
Pasul 3: $O(n)$
pasul 4: presupunând că cel puțin $\frac{3}{10}$ din tablou este $\leq p$ și că cel puțin $\frac{3}{10}$ din tablu este $\geq p$, pasul recursiv ia cel mult $T(\frac{7n}{10})$.
Însumând obținem:

$$
\begin{aligned}
T(n) &\leq cn + T(\frac{n}{5}) + T(\frac{7n}{10}) \\
&\leq cn + c\frac{n}{5} + T(\frac{n}{5^2}) + T(\frac{7n}{5 \cdot 10}) + c\frac{7n}{10} + T(\frac{7n}{5 \cdot 10}) + T(\frac{7^2 n}{10^2}) \\
&\leq \ldots \\
&= O(n)
\end{aligned}
$$

(similar teoremei de master).

Se arată că