## A. Szimultán maximum-minimum kiválasztás

hamarosan…

## B. Mediánkiválasztási algoritmusok alsókorlát elemzése

hamarosan…

## C. A $k$-adik elem kiválasztás véletlenített algoritmusa

**Feladat:** Adott $n$ darab különböző elem közül válasszuk ki a nagyság szerint $k$-adik legkisebb elemet. (Analóg módon megfogalmazható a $k$-adik legnagyobb elem keresésének problémája is.)

**Megoldás (absztrakt):** Tároljuk az elemeket egy halmazban! Válasszunk ki egy véletlen elemet, majd alakítsunk ki az $A$ illetve a $B$ halmazokat úgy, hogy előbbibe a kiválasztott elemnél kisebb, utóbbiba a nagyobb elemeket tesszük. Ekkor könnyen megállapítható az $A$ halmaz elemszáma alapján a kiválasztott elem nagyság szerinti helye (például ha $|A| = 6$, akkor a 7. elemet vettük ki). Ekkor három lehetőségünk van:

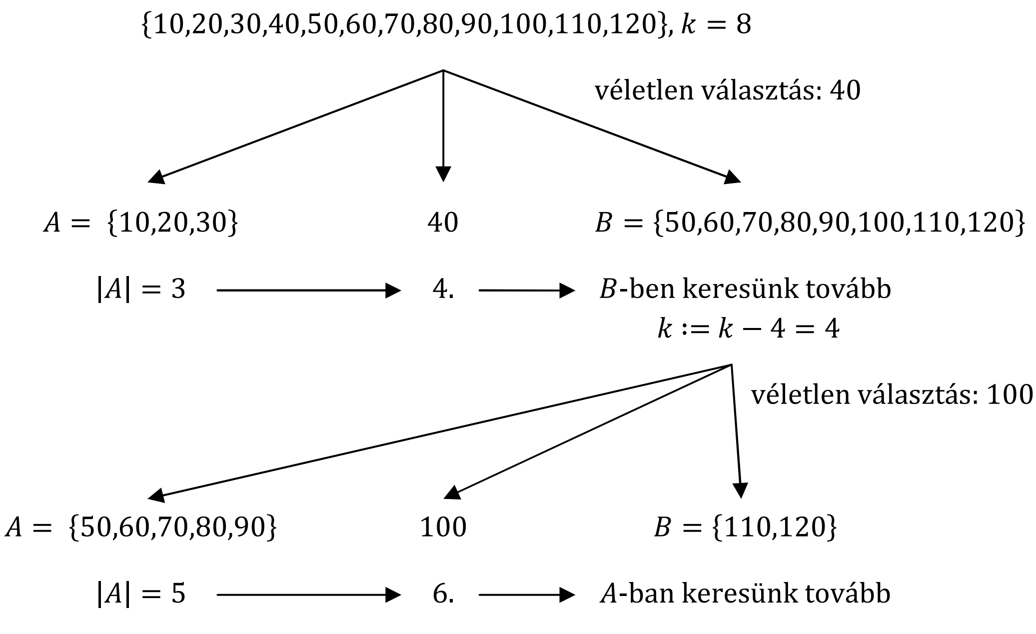
1) Ha $k$ megegyezik a kisorsolt elem helyével, akkor megtaláltuk a $k$-adik elemet, és leállunk.
2) Ha $k$ kisebb a véletlen elem helyénél, akkor az $A$ halmazra kell rekurzívan alkalmaznunk az algoritmust, azaz ott kell tovább keresnünk a $k$-adik elemet.
3) Ha $k$ nagyobb a kiválasztott elem helyénél, akkor a $B$ halmazra kell rekurzívan alkalmaznunk az algoritmust. Ebben az esetben módosítanunk kell a keresett $k$ értéket, hiszen a halmaz elemeit, és még egy további elemet már kivettük a keresési tartományból, így tehát $k := k - |A| - 1$.

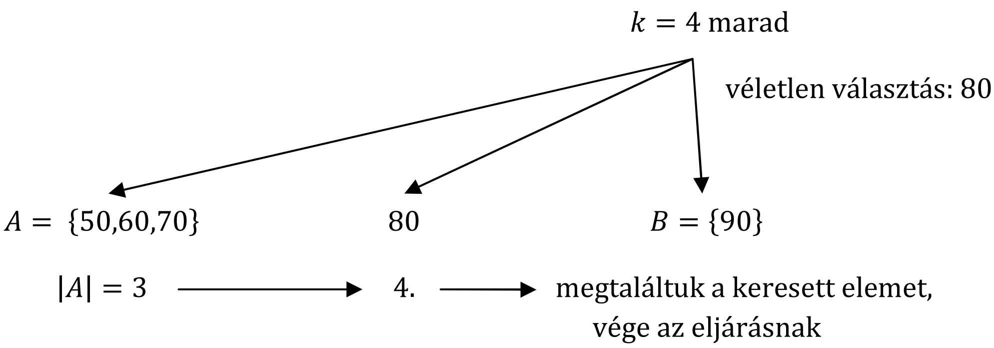
Mivel a keresés alatt álló halmaz elemszáma minden lépésben garantáltan legalább eggyel csökken, az eljárás véges időn belül terminál.

**Példa:** Tekintsük a $\{10,20,\dots,120\}$ halmazt, és keressük a $k = 8$. legkisebb elemet!

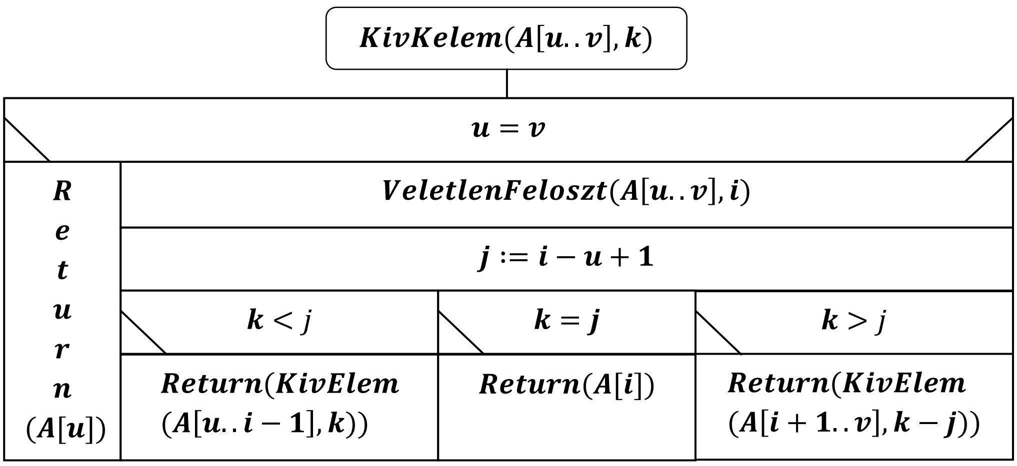
1) A véletlen választás eredménye legyen 40, így ennek megfelelően bontsuk szét a halmazt: $A = \{10,20,30\}, B = \{50,60,\dots,120\}$. Mivel $|A| = 3$, tudjuk, hogy a 40 a 4. elem a számsorozatban, ezért a 8. elemet a $B$ halmazban kell keresnünk. Mivel a számsorozat első 4 elemét leválasztottuk, ezért a $B$ halmazban már nem a 8., hanem a 4. elemet kell keresnünk.
2) A bemenő halmazunk $\{50,60,\dots,120\}$, a véletlen választás eredménye 100, így $A = \{50,\dots,90\}, B = \{110,120\}$ halmazok keletkeznek, így 100 a 6. legkisebb elem, tovább kell keresnünk a $A$ halmazban, továbbra is a 4. elemet.
3) A bemenő halmazunk $\{50,\dots,90\}$, a véletlen választás eredménye 80, így $A = \{50,60,70\}$, $B = \{90\}$, tehát a 80 a 4. elem, vagyis megtaláltuk a $k = 8$. legkisebb elemet.

Röviden illusztrálva az algoritmus működése:






**Megoldás (tömbös reprezentáció):** Adott tehát az $A[1..n]$ tömb, amely $n$ számú különböző értéket tartalmaz. Ebből szeretnénk kiválasztani azt az elemet, amelynek értéke $k$-adik az elemek nagyság szerint növekvő sorában.
Tegyük fel, hogy rendelkezünk egy olyan, az intervallumot véletlen módon felosztó algoritmussal, mint amilyet a gyorsrendezésnél használtunk „Helyrevisz” névvel. Ez – most valóban véletlenszerűen – kiválasztja az adott részintervallum egyik elemét, azt a nagyság szerinti helyére juttatja, és a nála kisebbeket tőle balra, a nála nagyobbakat tőle jobbra helyezi el.
A $VeletlenFeloszt$ eljárás a véletlen elem indexét adja vissza az i változóban, ahol tehát $u \leq i \leq v$. Ebből előállítjuk a véletlenül kiválasztott $A[i]$ elem nagyság szerint $j$ sorszámát: $j = i - u + 1$, mert mind a két értéket kifejező módon használjuk a maga helyén.
A $KivKelem(A[u..v], k)$ rekurzív algoritmus külső hívása a teljes $A[1..n]$ tömbbel és az eredetileg megadott $k$ értékkel történik. Később az eljárás önmagát egy résztömbre hívja meg, esetleg módosított k értékkel. Mindig fennáll azonban a $Q = (u \leq k \wedge 1 \leq k \leq v - u + 1)$ előfeltétel.




**Műveletigény:** Mivel véletlenkiválasztó algoritmust tárgyalunk, a műveletigényt nagyban befolyásolja egy elem kiválasztása a halmazból. Tekintsük az összehasonlítások számát az egyes esetekben. A legjobb, illetve legrosszabb eset műveletigénye független $k$értékétől, ezért itt azt elhanyagoljuk, csak az általános esetben vesszük figyelembe.

a) **Legjobb eset:** Elsőre sikerül kiválasztanunk a keresett elemet, ekkor csak a többi elemet kell – ezzel összehasonlítva – a megfelelő részhalmazba elhelyeznünk, ami $n - 1$ összehasonlítást jelent.

$$m\ddot{O}(n) = n - 1 = \Theta(n)$$

b) **Legrosszabb eset:** Egyrészt a véletlenválasztás mindig egy elemet választ le a többiről, ekkor az egyik halmaz üres, a másik pedig eggyel kisebb lesz, mint a korábbi halmaz. Másrészt egyszer sem emeli ki a keresett elemet, csak az utolsó lépésben. Ekkor az összehasonítások száma mindig eggyel csökken, míg egy elemű nem lesz a halmazunk.

$$M\ddot{O}(n) = (n-1) + (n-2) + \cdots + 2 + 1 = \frac{n(n-1)}{2} = \Theta(n^2)$$

c) **Átlagos esetben** igazolható, hogy az összehasonlítások száma lineáris műveletigényű, egész pontosan $A\ddot{O}(n) \leq 4n$. Vezessük be a következő segédfüggvényeket:
   1) $f(n,k) = A\ddot{O}(n,k)$, ahol $k$ a feladatban szereplő keresett nagyságrendi sorszám.
   2) $f_j(n,k)$ jelentse az átlagos műveletigényt, amikor az algoritmus a nagyság szerinti $j$-edik elemet választotta ki a halmazból. Nyilván annak függvényében, hogy $j$ hogy viszonyul $k$ értékéhez, három esetet különböztethetünk meg, és amennyiben nem egyenlők, akkor figyelembe kell vennünk a megfelelő rekurzív hívás műveletigényét az elemszám, illetve $k$ módosítása mellett, tehát:

$$f_j(n,k) = \begin{cases} n-1, ha\ j = k \\ (n-1) + f(j-1,k), ha\ k < j \\ (n-1) + f(n-j, k-j), ha\ k > j \end{cases}$$

Bizonyítsunk állításunkat teljes indukcióval:
   1) Kicsi n-re teljesül az állítás, például $f(1,1) \leq 4$, és $f(2,1) = f(2,2) \leq 8$
   2) Tegyük fel, hogy az állítás igaz $1,2,..n-1$ értékekre.
   3) Tekintsük $n$-re az átlagos esetet:

$$f(n,k) = \frac{1}{n}\sum_{j=1}^{n} f_j(n,k) = \frac{1}{n}n(n-1) + \frac{1}{n}\sum_{j=k+1}^{n} f(j-1,k) + \frac{1}{n}\sum_{j=1}^{k-1} f(n-j,k-j) \leq$$

$f_j(n,k)$ definíciója                                        indukciós feltevés

$$\leq \frac{1}{n}n(n-1) + \frac{1}{n}\sum_{j=k+1}^{n} 4(j-1) + \frac{1}{n}\sum_{j=1}^{k-1} 4(n-j) =$$

$$= (n-1) + \frac{1}{n}\sum_{j=k}^{n-1} 4j + \frac{1}{n}\sum_{j=1}^{k-1} 4(n-j) = (n-1) + \frac{4}{n}\sum_{j=1}^{n-1} j - \frac{4}{n}\sum_{j=1}^{k-1} j + \frac{4}{n}\sum_{j=1}^{k-1}(n-j) =$$

$$= (n-1) + \frac{4}{n}\frac{n(n-1)}{2} + \frac{4}{n}\sum_{j=1}^{k-1}(n-2j) =$$

$$= 3(n-1) + \frac{4}{n}(k-1)\frac{(n-2) + \big(n-2(k-1)\big)}{2} =$$

$$= 3(n-1) + \frac{4}{n}(k-1)(n-k)\frac{2n-2k}{2} = 3(n-1) + \frac{4}{n}\left(\frac{n-1}{2}\right)^2 \leq 3n + \frac{4}{n}\cdot\frac{n^2}{4} = 4n$$

Tehát beláttuk, hogy tetszőleges $n, k$ értékekre $f(n,k) \leq 4$, az eredeti állítás fennáll.

$$A\ddot{O}(n) \leq 4n$$