Algorithm: Sel(numbers,k)

Given: numbers $x_1,\ldots,x_n$, k, output k'th smallest number.

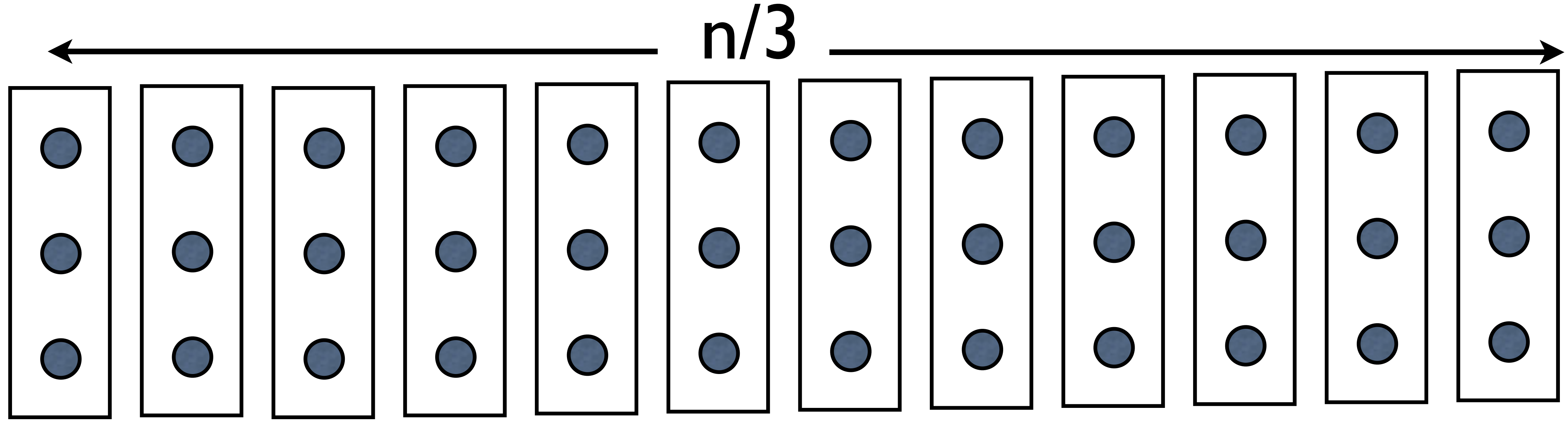
Algorithm: Sel(numbers,k)

Given: numbers $x_1,\ldots,x_n$, k, output k'th smallest number.

Sort the numbers!   O(n log n) time.

Can we do better?

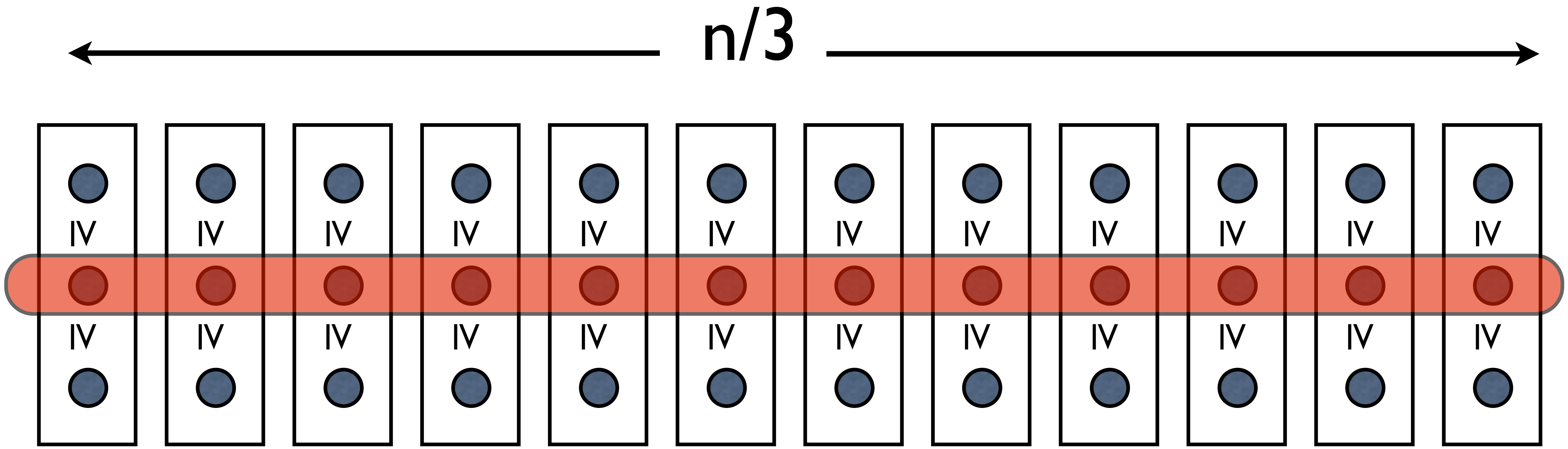
Given: $x_1,\ldots,x_n$, k, output k'th smallest number.



Algorithm: Sel(numbers,k)

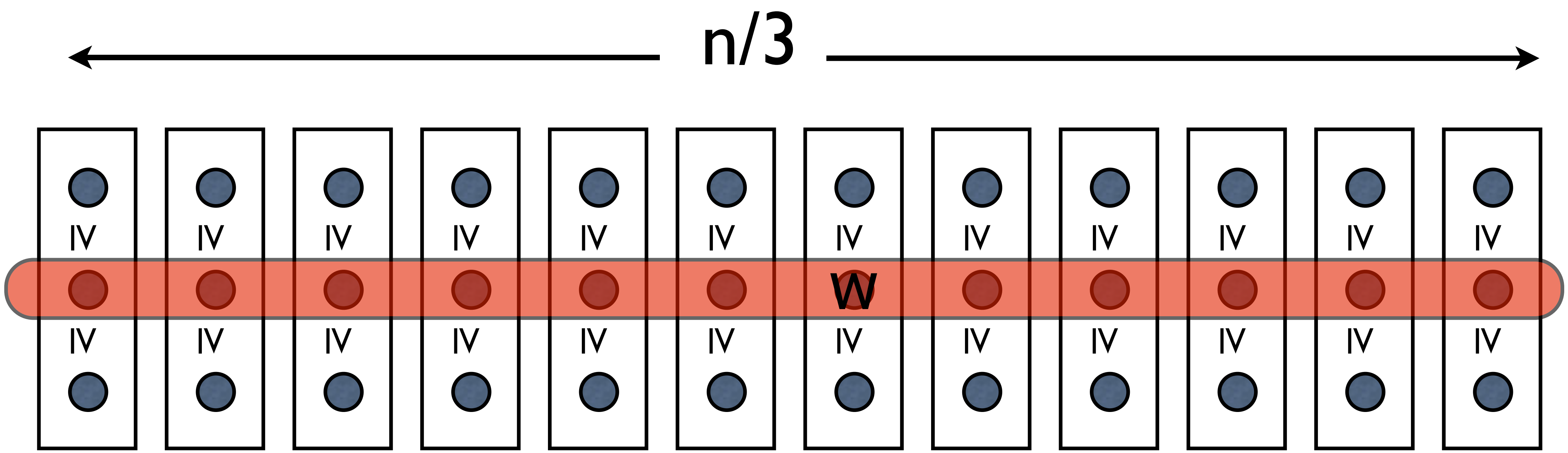
1. Partition numbers into sets of size 3

Algorithm: Sel(numbers,k)



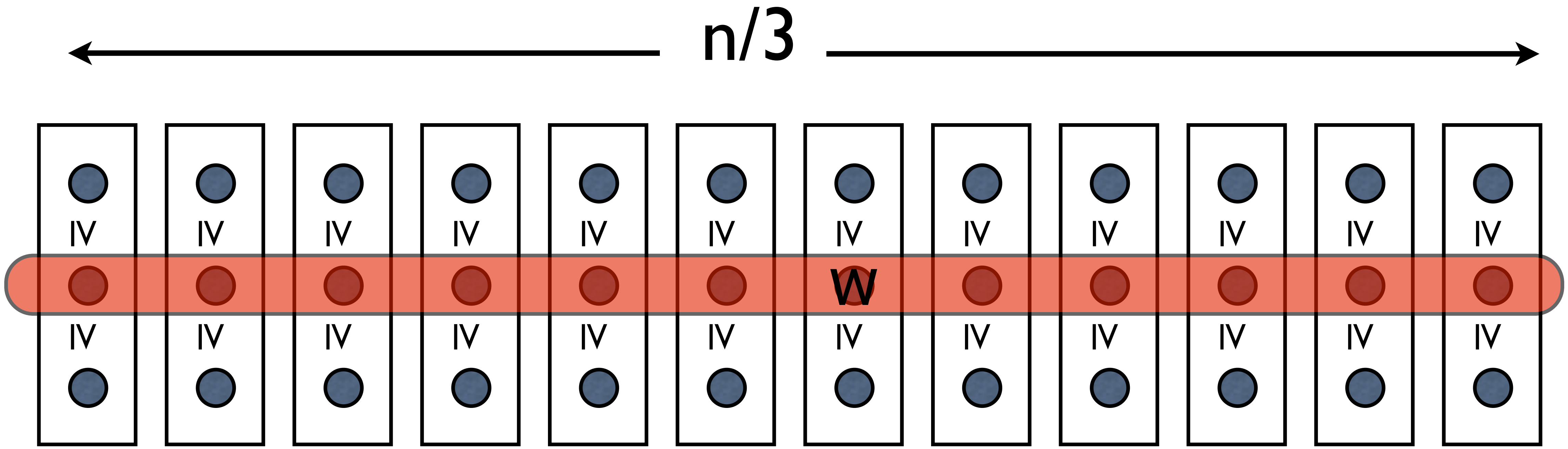
1. Partition numbers into sets of size 3
2. Sort each set

Algorithm: Sel(numbers,k)



1. Partition numbers into sets of size 3
2. Sort each set
3. w = Sel(,n/6)

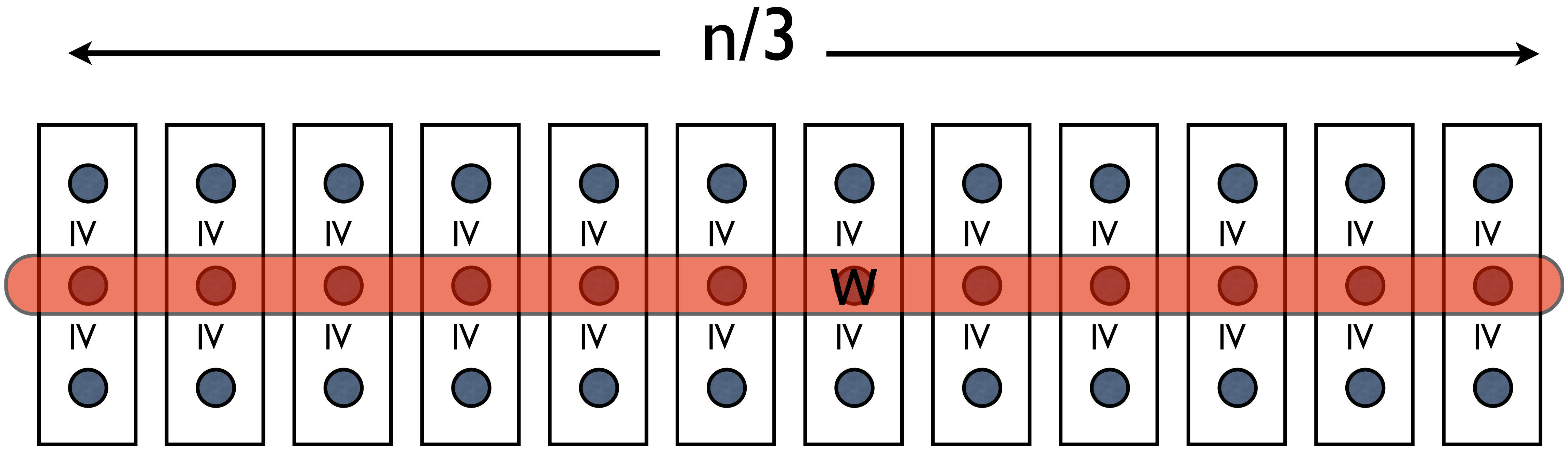
Algorithm: Sel(numbers,k)



1. Partition numbers into sets of size 3
2. Sort each set
3. w = Sel(,n/6)

$$\left.\begin{aligned} S_L(w) &= \{x_i \mid x_i < w\} \\ S_E(w) &= \{x_i \mid x_i = w\} \\ S_G(w) &= \{x_i \mid x_i > w\} \end{aligned}\right\}$$ Can be computed in linear time

Algorithm: Sel(numbers,k)

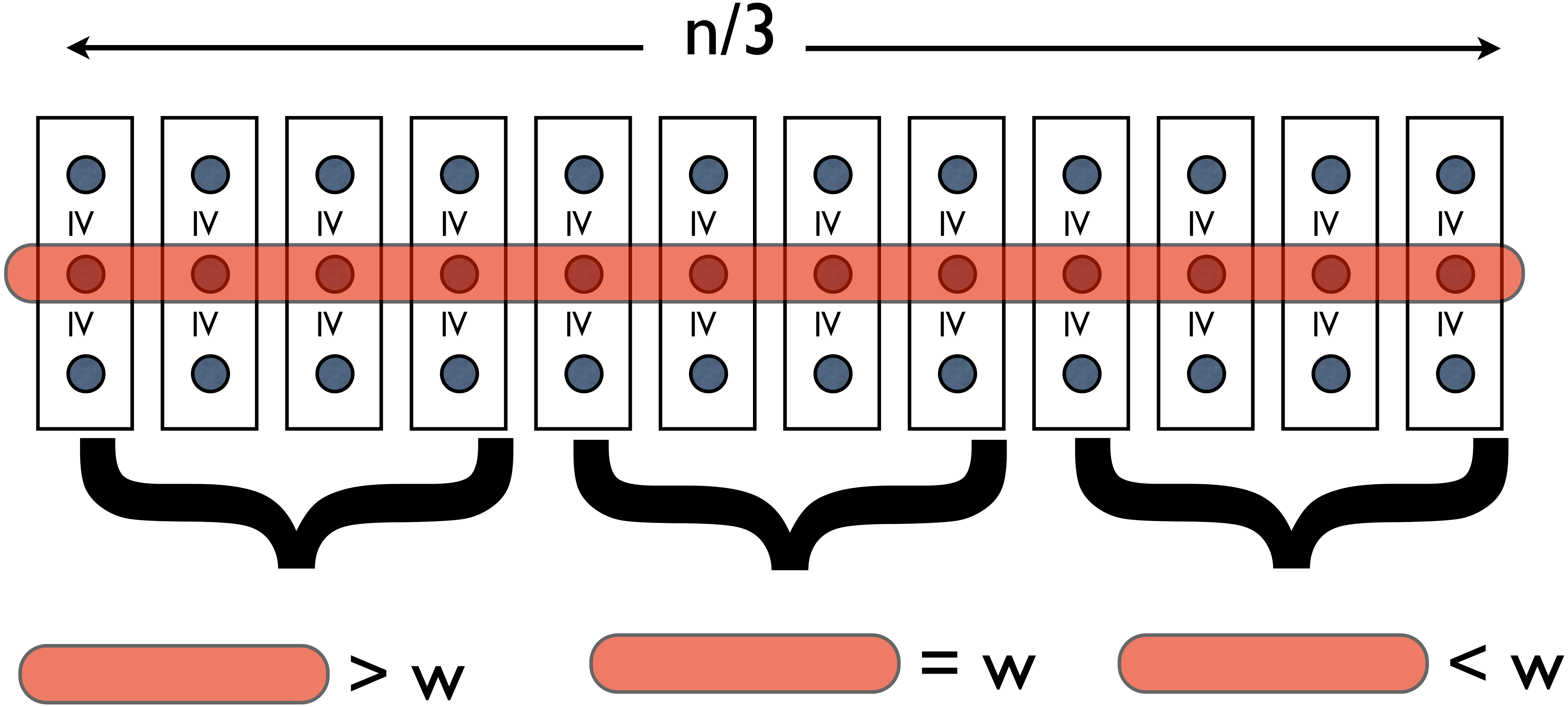


1. Partition numbers into sets of size 3
2. Sort each set
3. w = Sel(,n/6)

$$\left.\begin{array}{l} S_L(w) = \{x_i \mid x_i < w\} \\ S_E(w) = \{x_i \mid x_i = w\} \\ S_G(w) = \{x_i \mid x_i > w\} \end{array}\right\}$$ Can be computed in linear time

4. if $k \le |S_L(w)|$, output $Sel(S_L(w), k)$
   else if $k \le |S_L(w)| + |S_E(w)|$, output w
   else output $Sel(S_G(w), k - |S_L(w)| - |S_E(w)|)$

Algorithm: Sel(numbers,k)



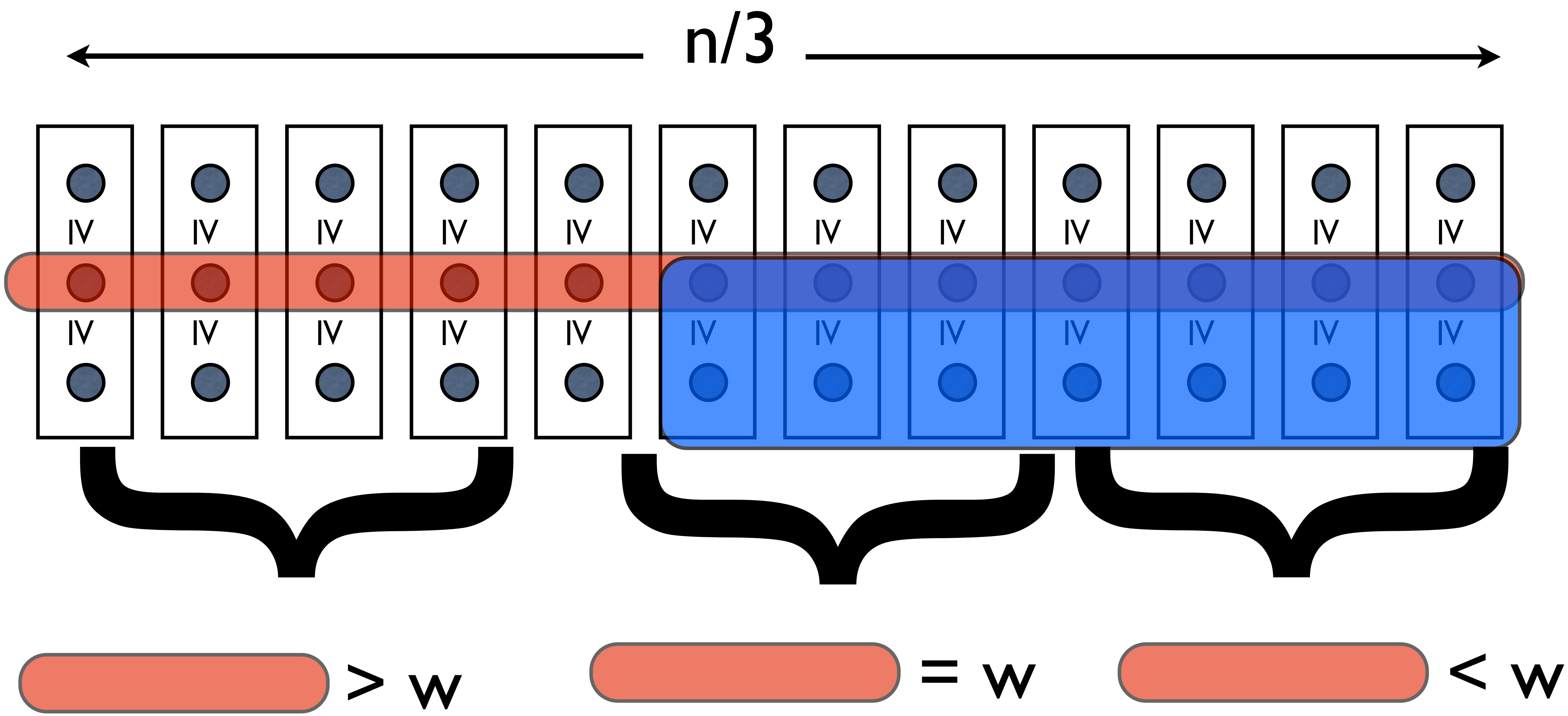
w = Sel( ,n/6)

if $k \le |S_L(w)|$, output Sel($S_L(w)$, k)

else if $k \le |S_L(w)| + |SE(w)|$, output w

else output Sel($S_G(w)$, k - $|S_L(w)|$-$|S_G(w)|$)

Algorithm: Sel(numbers,k)



$|S_L(w)| + |S_E(w)|$, at least $2(n/6) = n/3$

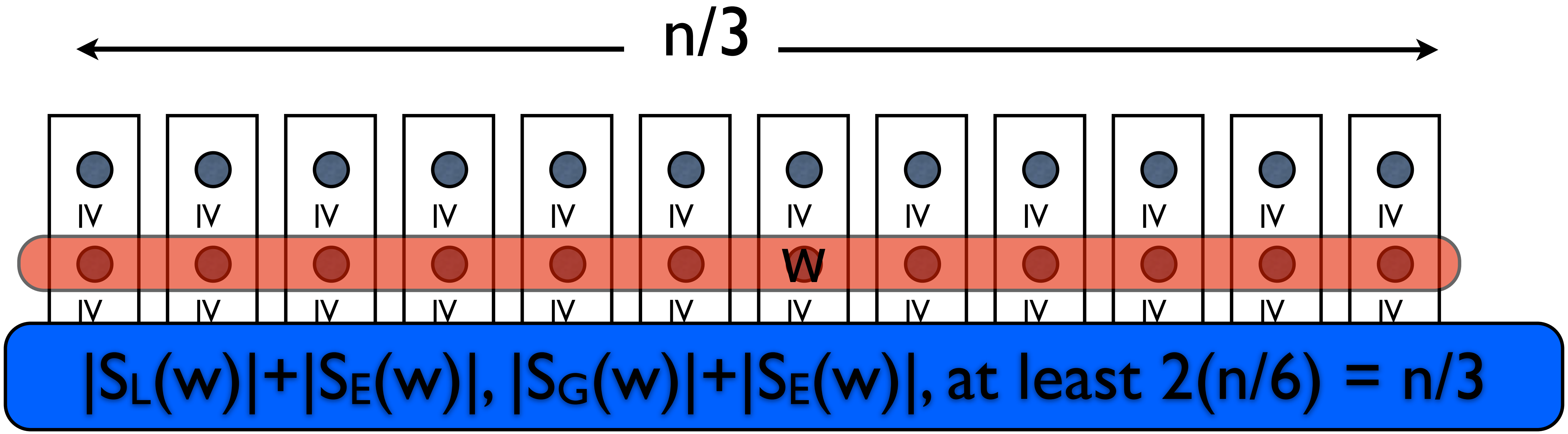
w = Sel(,n/6)

if $k \leq |S_L(w)|$, output Sel($S_L(w)$, k)

else if $k \leq |S_L(w)| + |SE(w)|$, output w

else output Sel($S_G(w)$, k - $|S_L(w)|$-$|S_G(w)|$)

Algorithm: Sel(numbers,k)



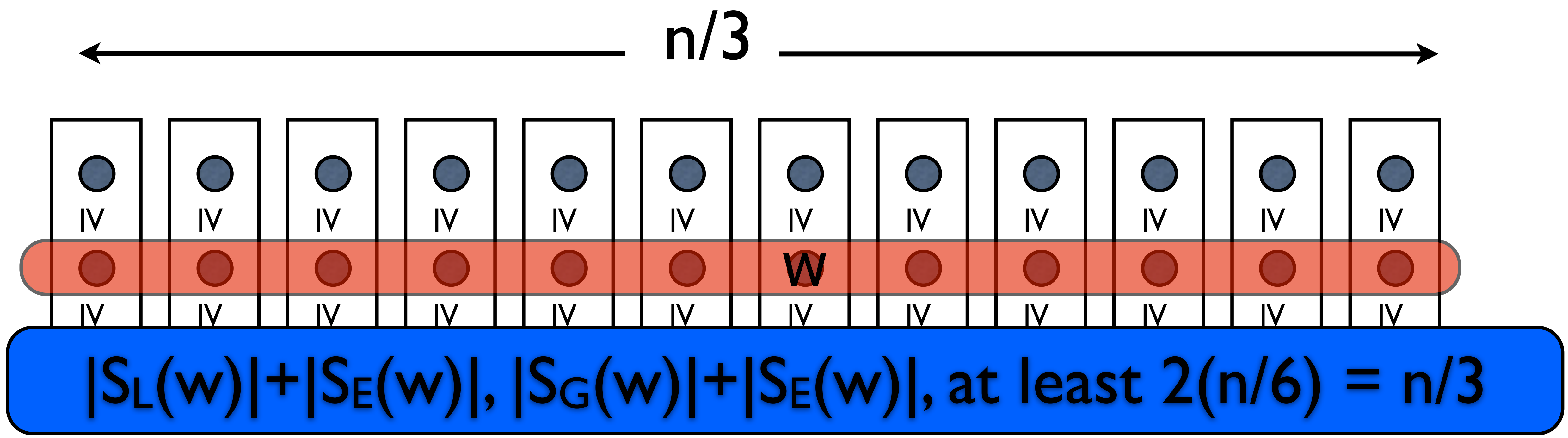
1. Partition numbers into sets of size 3
2. Sort each set
3. w = median of

$$T(n) = T(n/3) + T(2n/3) + O(n)$$

$$\left.\begin{array}{l} S_L(w) = \{x_i \mid x_i < w\} \\ S_E(w) = \{x_i \mid x_i = w\} \\ S_G(w) = \{x_i \mid x_i > w\} \end{array}\right\} \text{Can be computed in linear time}$$

4. if $k \le |S_L(w)|$, output $Sel(S_L(w), k)$
   else if $k \le |S_L(w)| + |S_E(w)|$, output w
   else output $Sel(S_G(w), k - |S_L(w)| - |S_E(w)|)$

Algorithm: Sel(numbers,k)



1. Partition numbers into sets of size 3
2. Sort each se
3. w = median

$S_L(w) = \{x_i$

$S_E(w) = \{x_i$

$S_G(w) = \{x_i \mid x_i > w\}$

linear time

$T(n) = T(n/3) + T(2n/3) + O(n)$

so

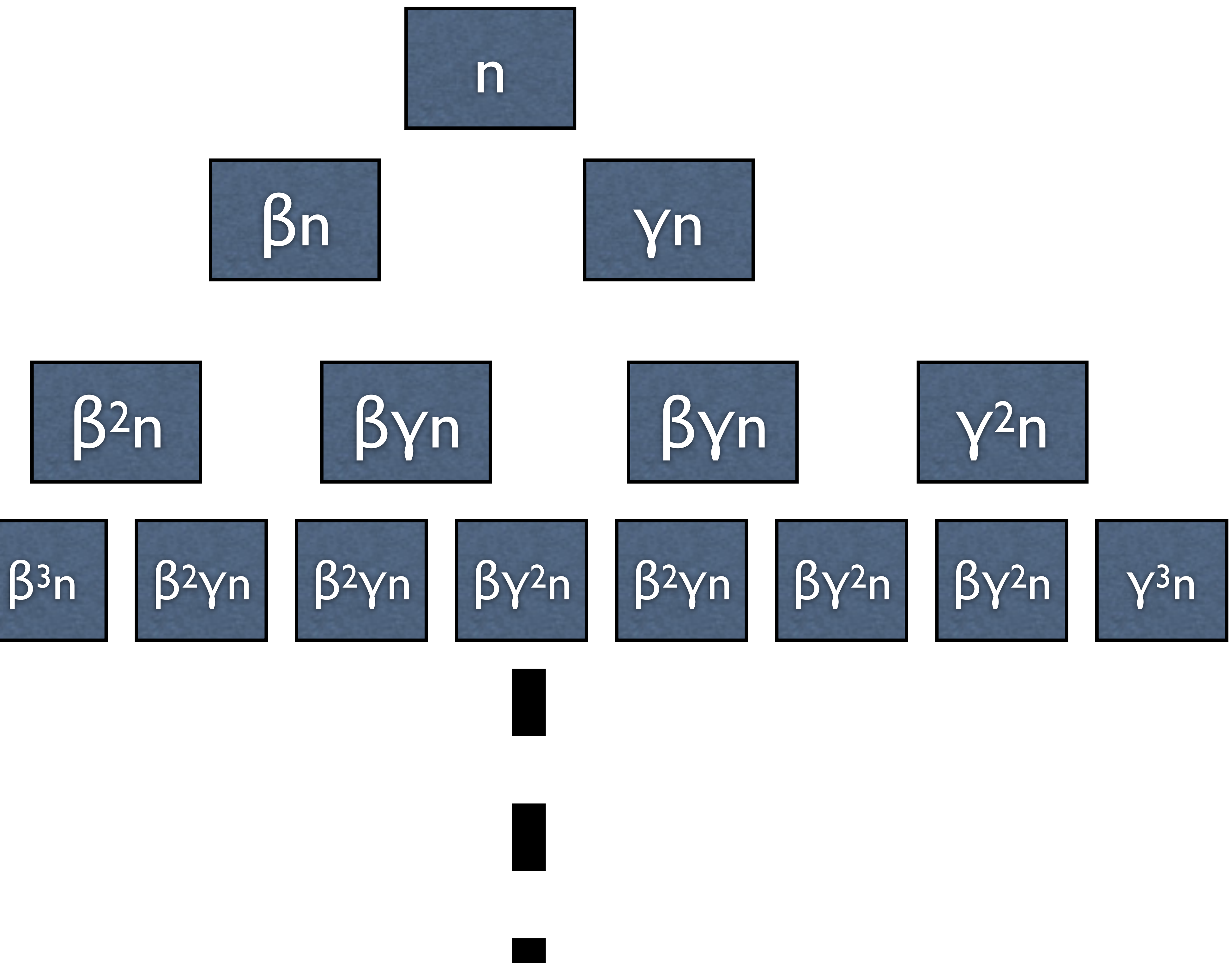
$T(n) = O(n\log n)$

(what's the point???)

4. if $k \leq |S_L(w)|$, output $Sel(S_L(w), k)$
   else if $k \leq |S_L(w)| + |S_E(w)|$, output w
   else output $Sel(S_G(w), k - |S_L(w)| - |S_E(w)|)$

Recurrences

$$T(n) = T(\gamma n) + T(\beta n) + n$$



n

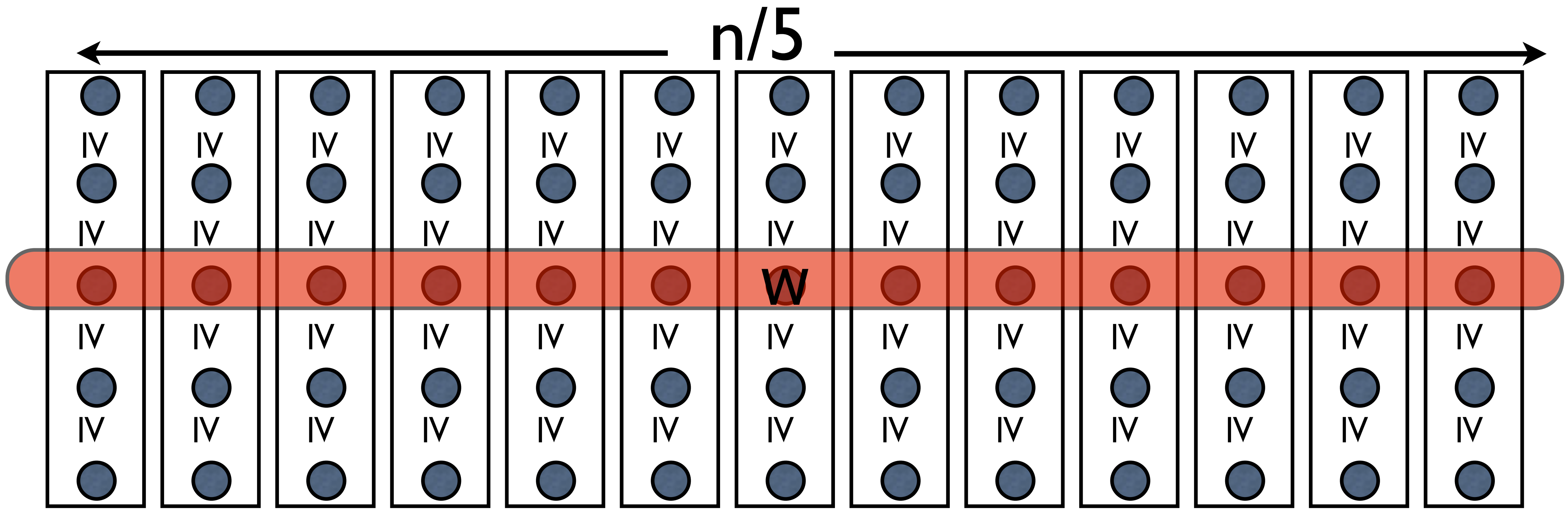
$(\beta+\gamma)n$

$(\beta+\gamma)^2 n$

$(\beta+\gamma)^3 n$

$$n \sum_i (\beta+\gamma)^i$$

Algorithm: Sel(numbers,k)



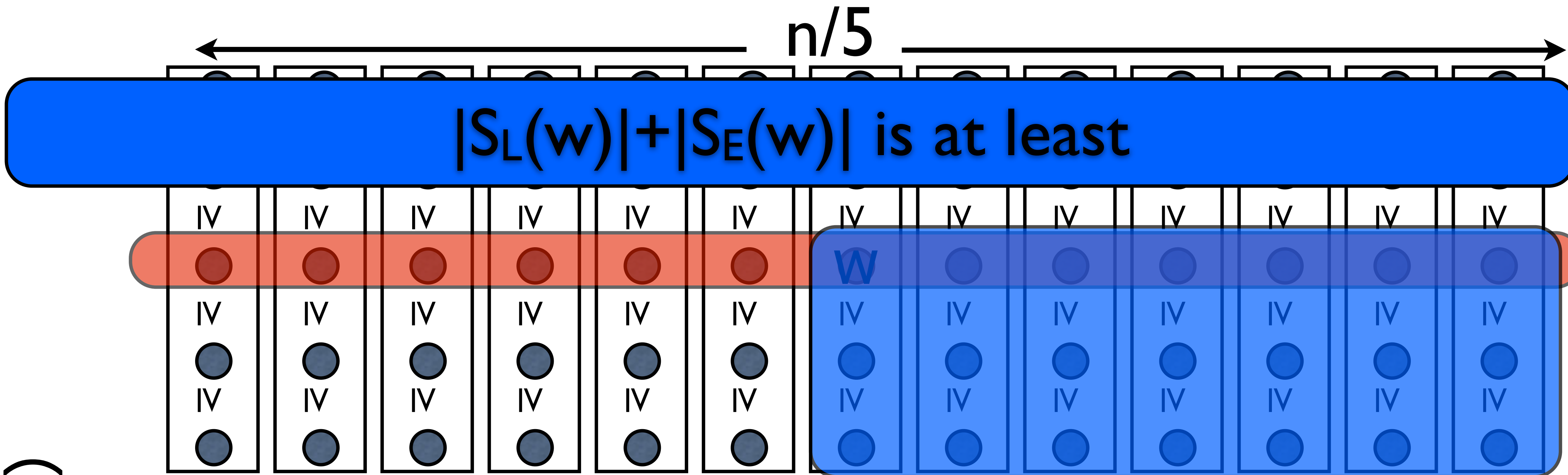
1. Partition numbers into sets of size 5
2. Sort each set
3. w = median of

$$S_L(w) = \{x_i \mid x_i < w\}$$
$$S_E(w) = \{x_i \mid x_i = w\}$$
$$S_G(w) = \{x_i \mid x_i > w\}$$

Can be computed in linear time

4. if $k \le |S_L(w)|$, output $Sel(S_L(w), k)$
   else if $k \le |S_L(w)| + |S_E(w)|$, output w
   else output $Sel(S_G(w), k - |S_L(w)| - |S_E(w)|)$

Algorithm: Sel(numbers,k)



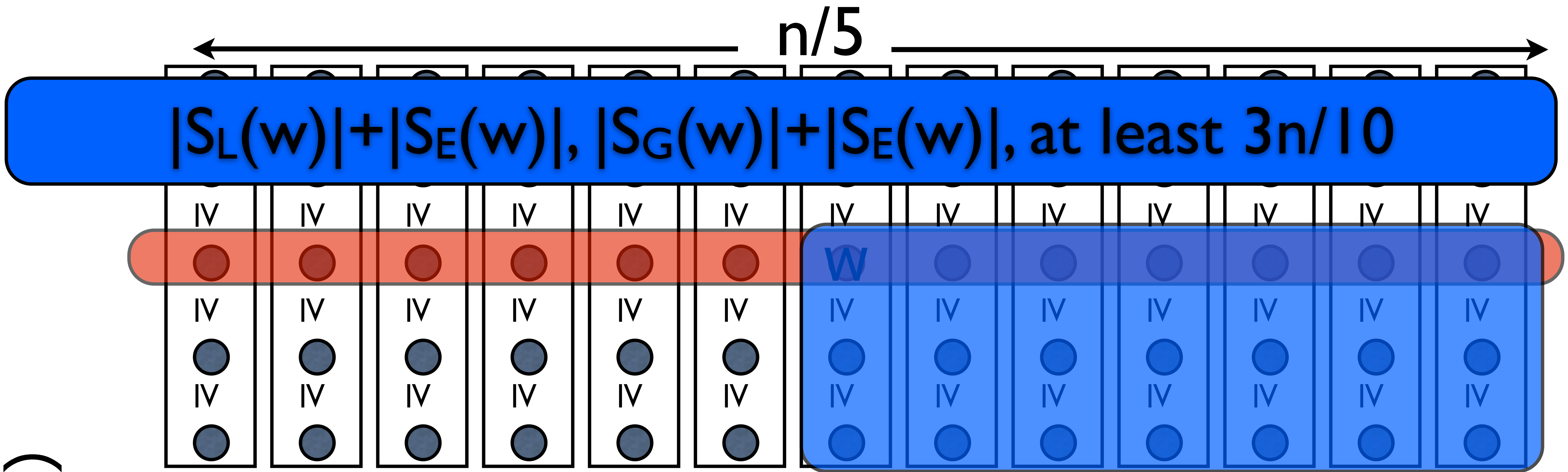
1. Partition numbers into sets of size 5
2. Sort each set
3. w = median of

$S_L(w) = \{x_i \mid x_i < w\}$
$S_E(w) = \{x_i \mid x_i = w\}$
$S_G(w) = \{x_i \mid x_i > w\}$

Can be computed in linear time

4. if $k \leq |S_L(w)|$, output $\mathrm{Sel}(S_L(w), k)$
   else if $k \leq |S_L(w)| + |S_E(w)|$, output w
   else output $\mathrm{Sel}(S_G(w), k - |S_L(w)| - |S_E(w)|)$

Algorithm: Sel(numbers,k)

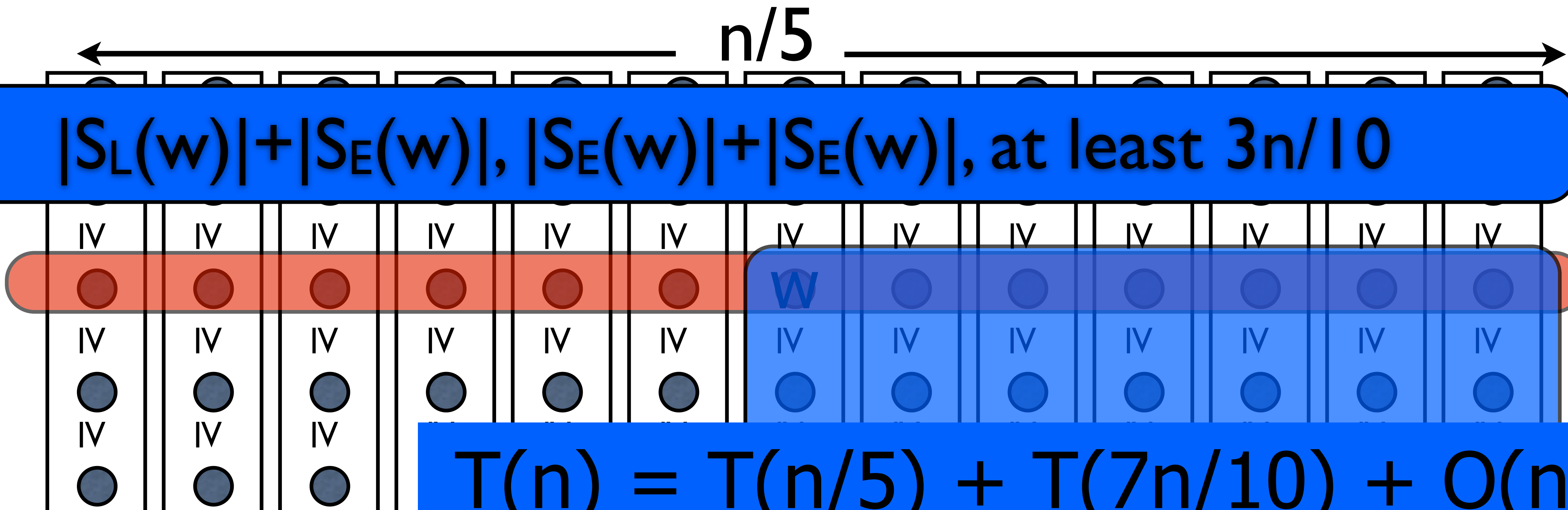


1. Partition numbers into sets of size 5
2. Sort each set
3. w = median of

$$\left.\begin{array}{l} S_L(w) = \{x_i \mid x_i < w\} \\ S_E(w) = \{x_i \mid x_i = w\} \\ S_G(w) = \{x_i \mid x_i > w\} \end{array}\right\} \text{Can be computed in linear time}$$

4. if $k \leq |S_L(w)|$, output Sel($S_L(w)$, k)
   else if $k \leq |S_L(w)| + |S_E(w)|$, output w
   else output Sel($S_G(w)$, k - $|S_L(w)|$-$|S_E(w)|$)

Algorithm: Sel(numbers,k)



$|S_L(w)|+|S_E(w)|$, $|S_E(w)|+|S_E(w)|$, at least $3n/10$

1. Partition nu
2. Sort each s
3. w = mediar

$$T(n) = T(n/5) + T(7n/10) + O(n)$$

so

$$T(n) = O(n)$$

$$\left.\begin{aligned} S_L(w) &= \{x_i \mid x_i < w\} \\ S_E(w) &= \{x_i \mid x_i = w\} \\ S_G(w) &= \{x_i \mid x_i > w\} \end{aligned}\right\}$$ Can be computed in linear time

4. if $k \leq |S_L(w)|$, output $Sel(S_L(w), k)$
   else if $k \leq |S_L(w)| + |S_E(w)|$, output w
   else output $Sel(S_G(w), k - |S_L(w)| - |S_E(w)|)$