

# DATA STRUCTURES & ALGORITHMS

## 20: MERGE SORT (DIVIDE & CONQUER)

Dr Ram Prasad Krishnamoorthy

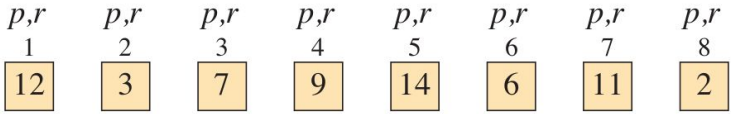
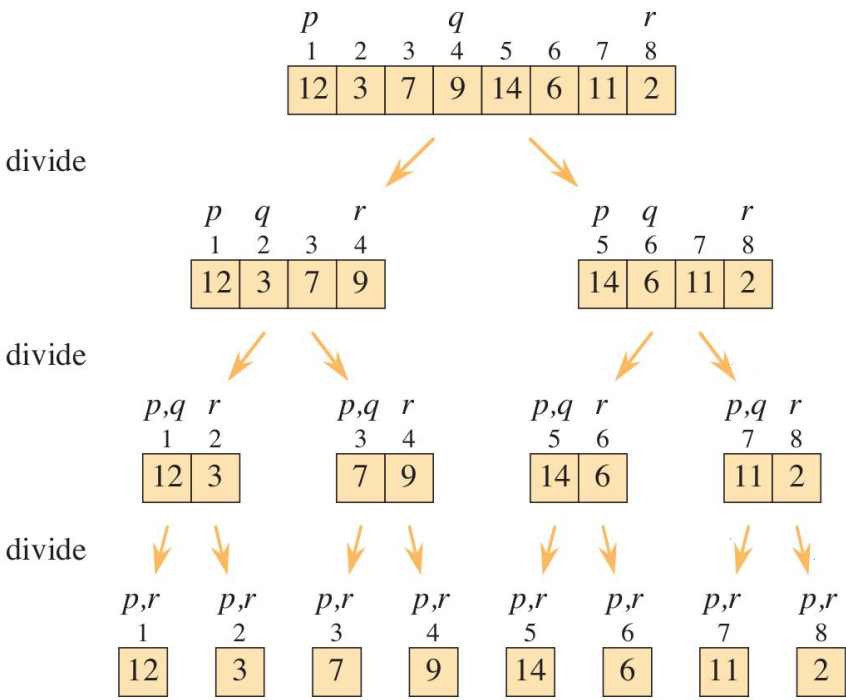
*Associate Professor  
School of Computing and Data Science*

[ram.krish@saiuniversity.edu.in](mailto:ram.krish@saiuniversity.edu.in)



# MERGE SORT

# MERGE SORT




SPLITTING IS DONE  
RECURSIVELY

# MERGE SORT

$p$			$q$			$r$		
1	2	3	4	5	6	7	8	
12	3	7	9	14	6	11	2	

divide

$l$  

$p$		$q$	$r$	
1	2	3	4	
12	3	7	9	

$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$
1	2	3	4	5	6	7	8
12	3	7	9	14	6	11	2

# MERGE SORT

$p$				$q$				$r$	
1	2	3	4	5	6	7	8		
12	3	7	9	14	6	11	2		

divide



$p$		$q$	$r$
1	2	3	4
12	3	7	9

divide



$p,q$	$r$
1	2
12	3

$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$
1	2	3	4	5	6	7	8
12	3	7	9	14	6	11	2

# MERGE SORT

$p$			$q$	$r$			
1	2	3	4	5	6	7	8
12	3	7	9	14	6	11	2

divide



$p$	$q$	$r$	
1	2	3	4
12	3	7	9

divide



$p,q$	$r$
1	2
12	3

divide



$p,r$	$p,r$
1	2
12	3

$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$
1	2	3	4	5	6	7	8
12	3	7	9	14	6	11	2

# MERGE SORT

$p$			$q$	$r$			
1	2	3	4	5	6	7	8
12	3	7	9	14	6	11	2

divide



$p$	$q$	$r$	
1	2	3	4
12	3	7	9

divide



$p,q$	$r$
1	2
12	3

divide

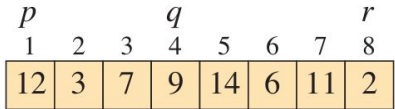


$p,r$	$p,r$
1	2
12	3

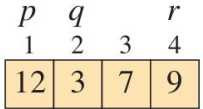
$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$	$p,r$
1	2	3	4	5	6	7	8
12	3	7	9	14	6	11	2



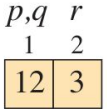
# MERGE SORT



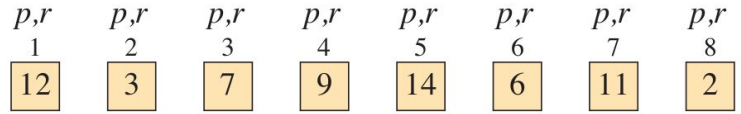
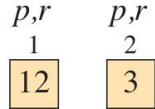
divide  $1$  




divide  $2$  



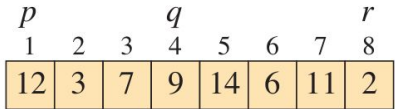
divide  $3$    $4$  



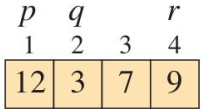
merge   

$p,q$	$r$
1	2
3	12

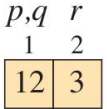
# MERGE SORT



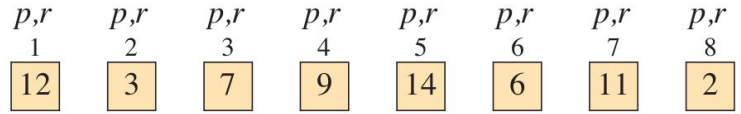
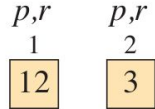
divide  $1$  



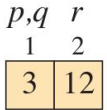
divide  $2$  



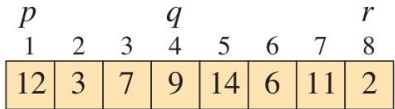
divide  $3$    $4$  



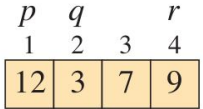
merge  $5$  



# MERGE SORT



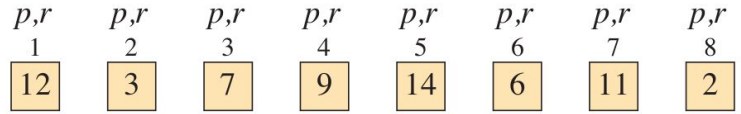
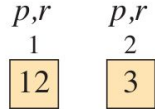
divide  $1$  



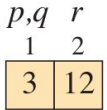
divide  $2$   



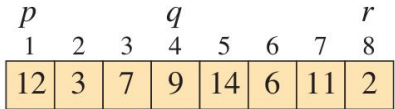
divide  $3$    $4$  



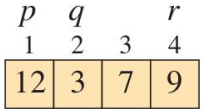
merge  $5$   



# MERGE SORT



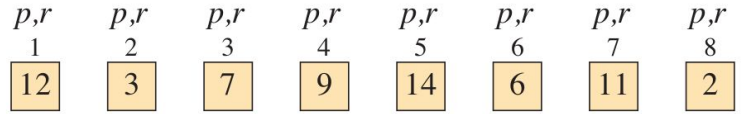
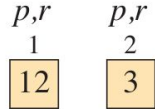
divide  $1$  ↘



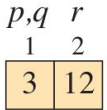
divide  $2$  ↘  $6$  ↘



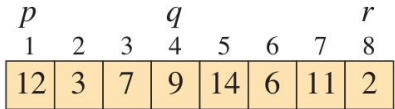
divide  $3$  ↘  $4$  ↘



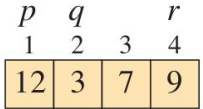
merge  $5$  ↘



# MERGE SORT



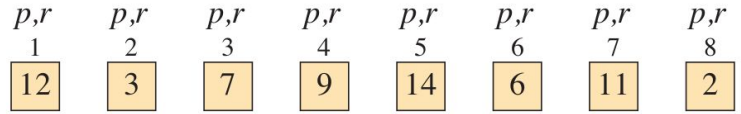
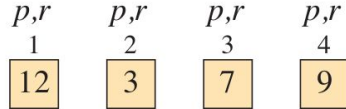
divide  $1$  ↘



divide  $2$  ↘  $6$  ↘



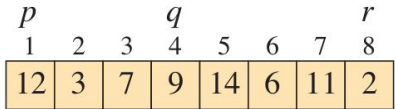
divide  $3$  ↘  $4$  ↘  $7$  ↘  $9$  ↘



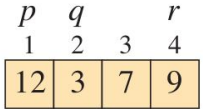
merge  $5$  ↘  $7$  ↘

$p,q$	$r$
1	2
3	12

# MERGE SORT



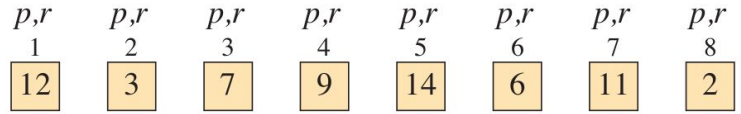
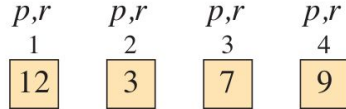
divide 1 ↘



divide 2 ↘ 6 ↘



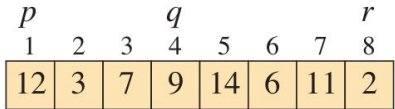
divide 3 ↘ 4 ↘ 7 ↘ 8 ↘



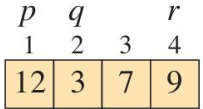
merge 5 ↙ 5 ↘

<i>p,q</i>	<i>r</i>
1	2
3	12

# MERGE SORT



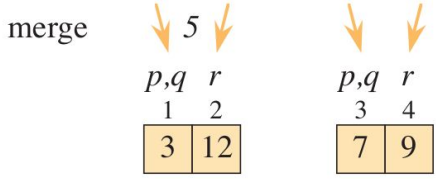
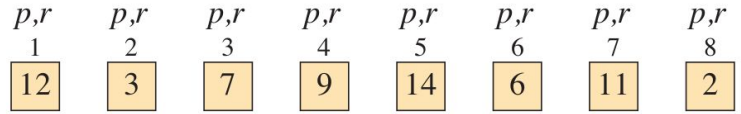
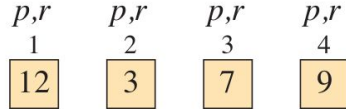
divide 1 ↘



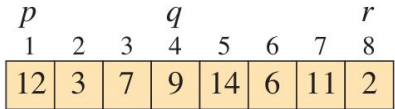
divide 2 ↘ 6 ↘



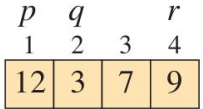
divide 3 ↘ 4 ↘ 7 ↘ 8 ↘



# MERGE SORT



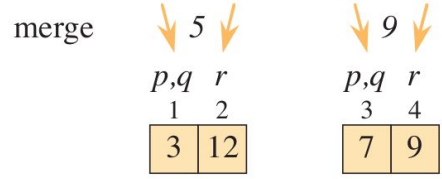
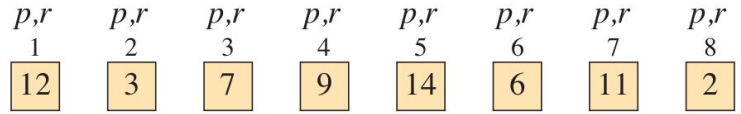
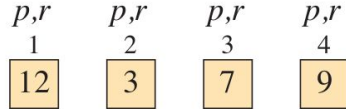
divide  $1 \searrow$



divide  $2 \searrow$   $6 \searrow$

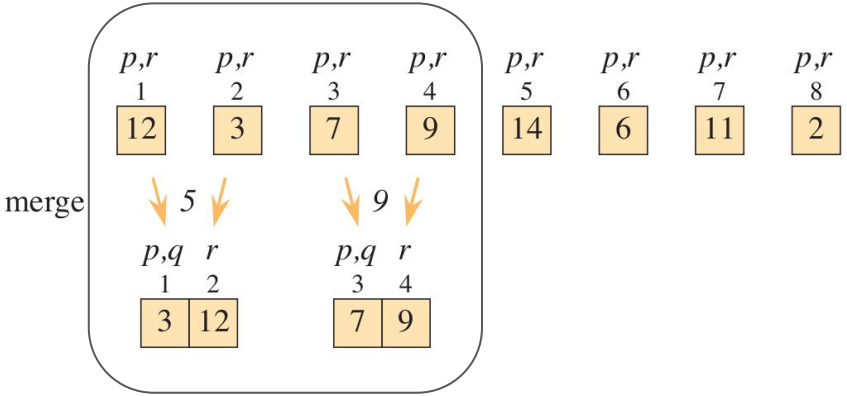
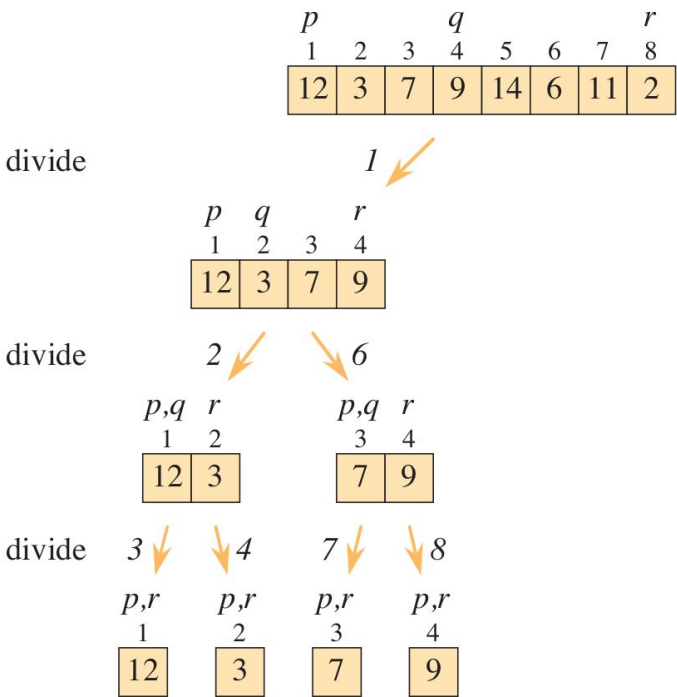


divide  $3 \searrow$   $4 \searrow$   $7 \searrow$   $8 \searrow$

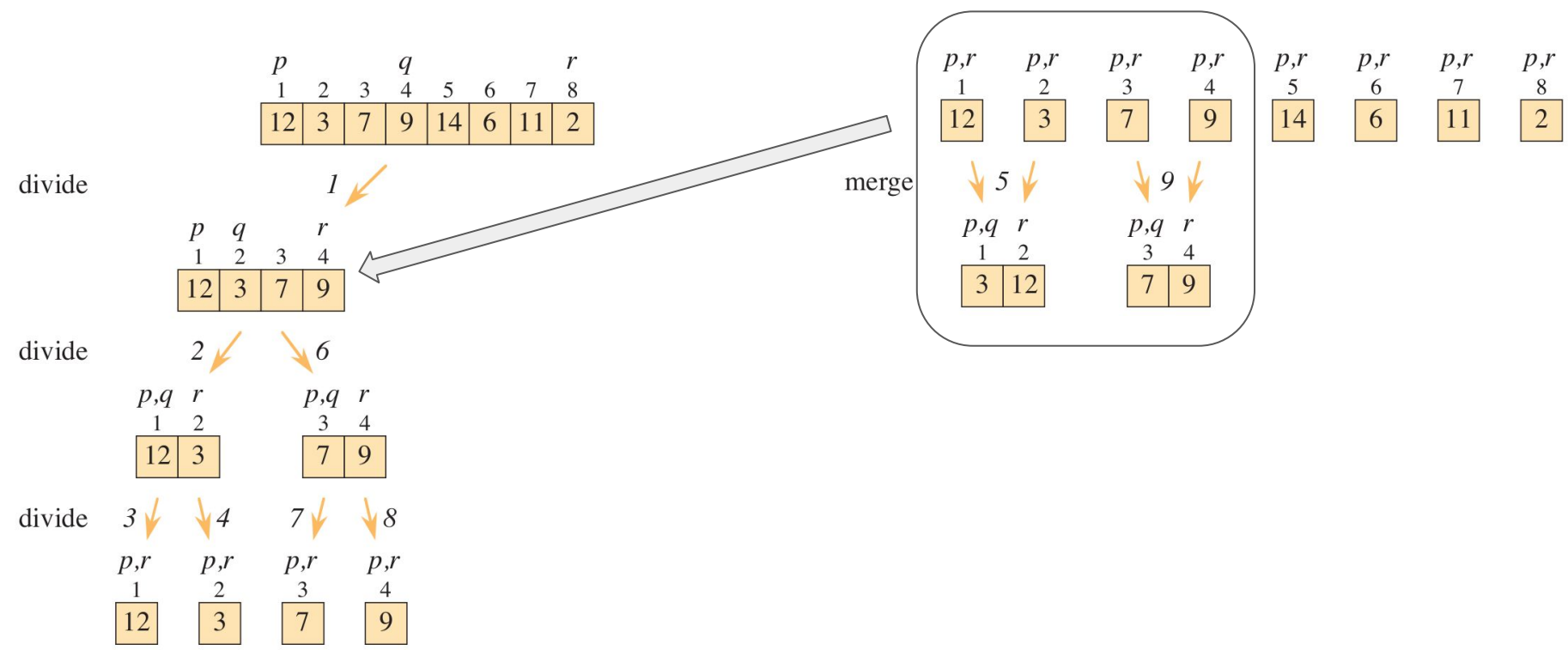




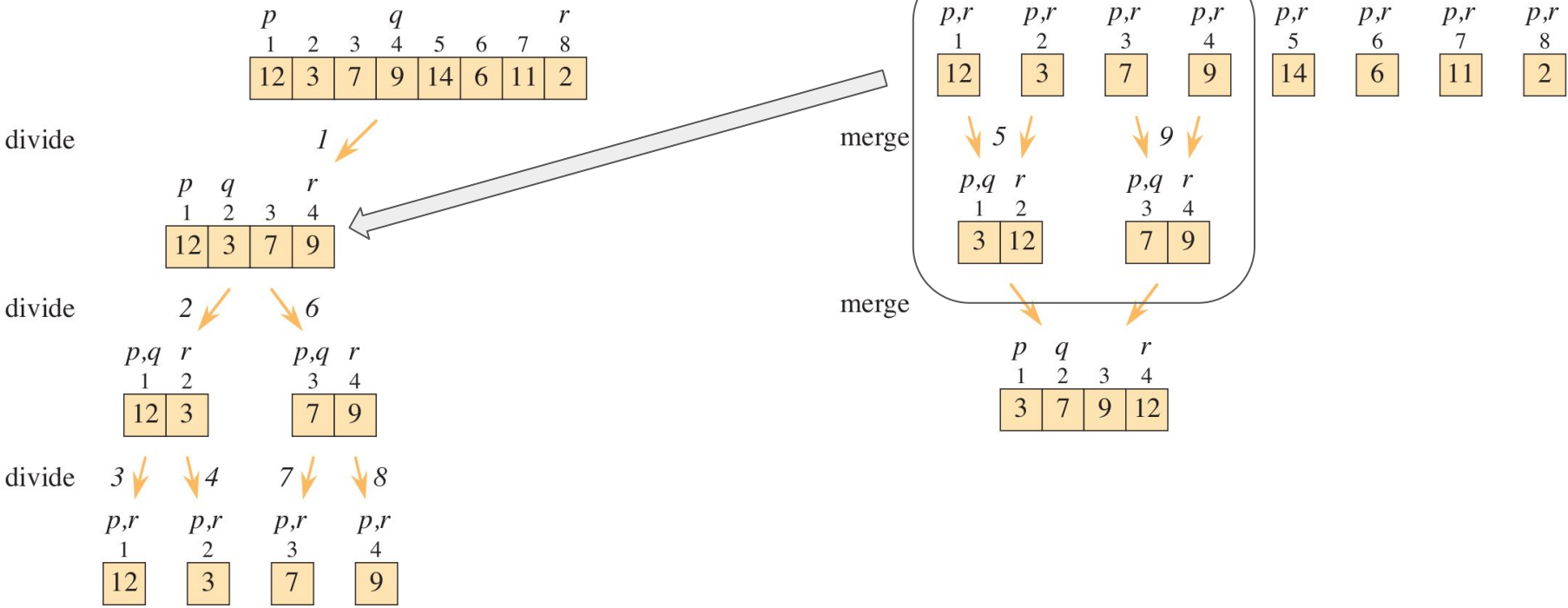
# MERGE SORT



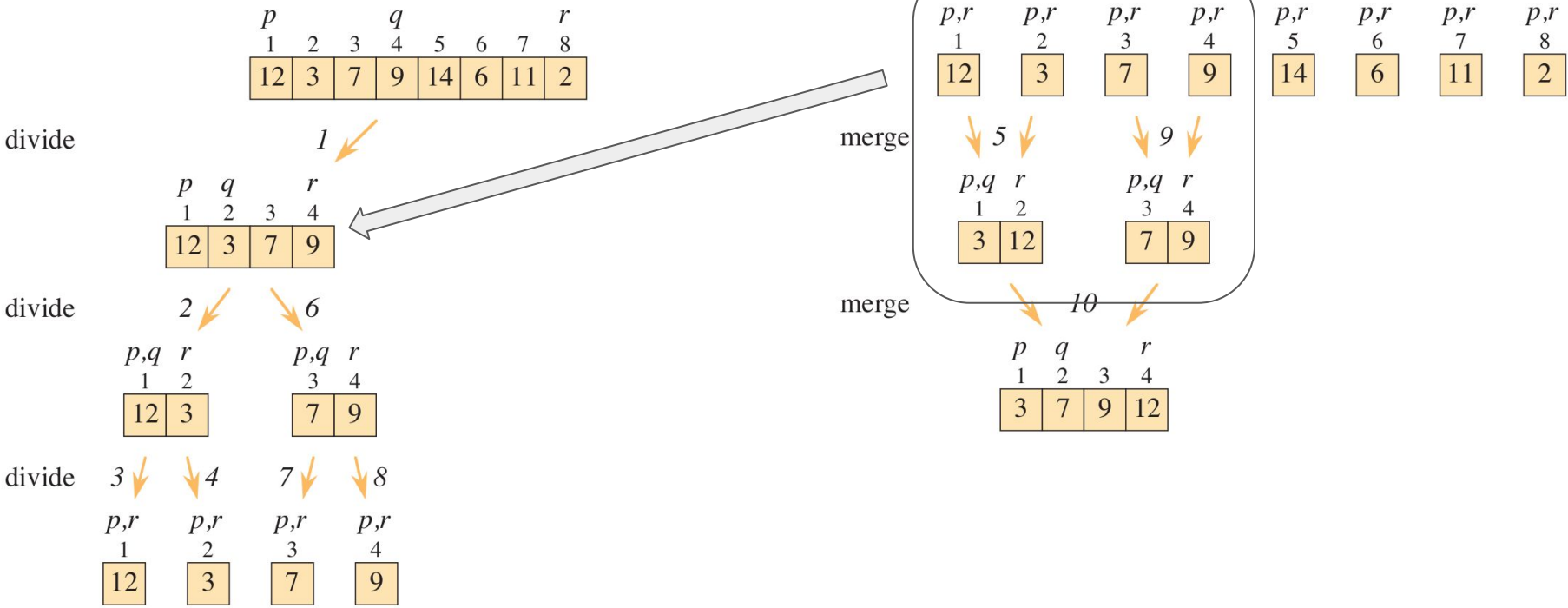
# MERGE SORT



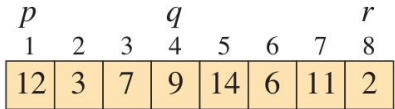
# MERGE SORT



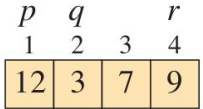
# MERGE SORT



# MERGE SORT



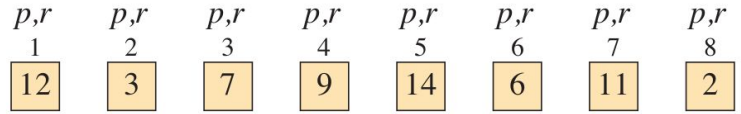
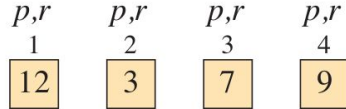
divide  $1 \searrow$



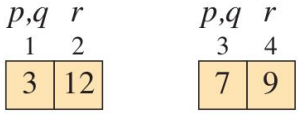
divide  $2 \searrow$   $6 \searrow$



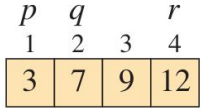
divide  $3 \searrow$   $4 \searrow$   $7 \searrow$   $8 \searrow$



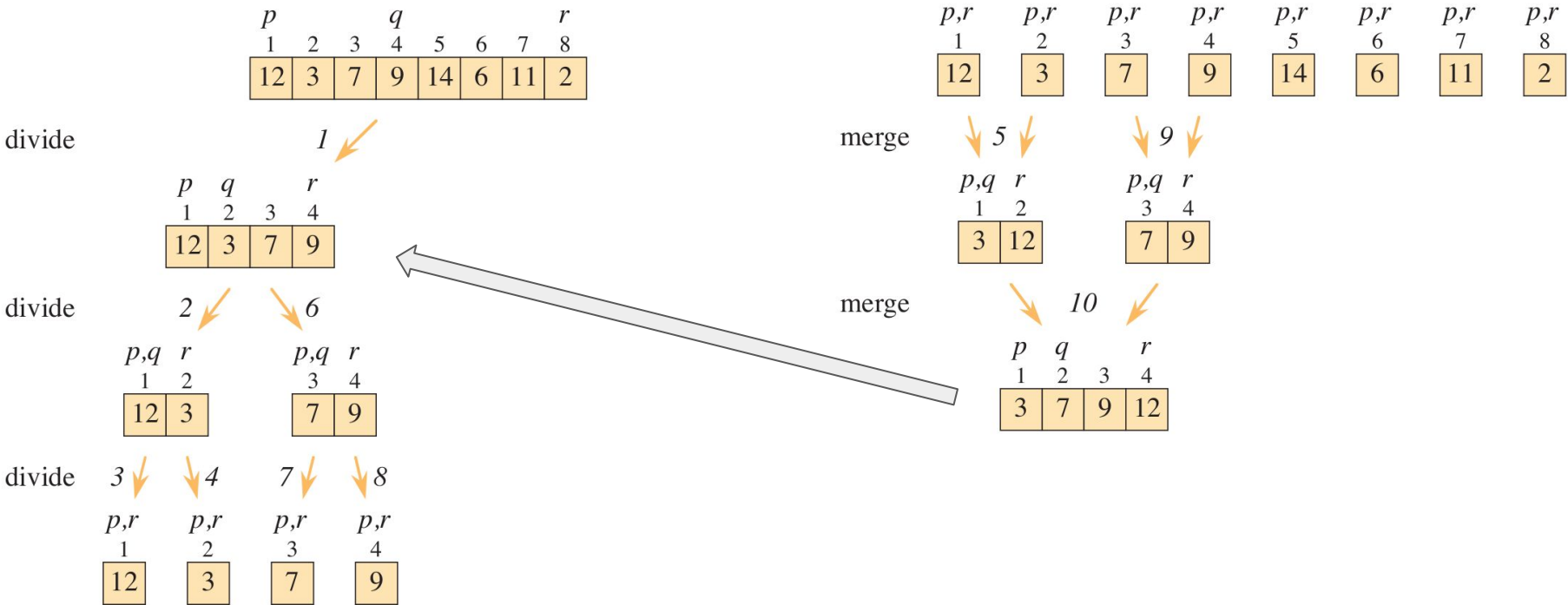
merge  $5 \searrow$   $9 \searrow$



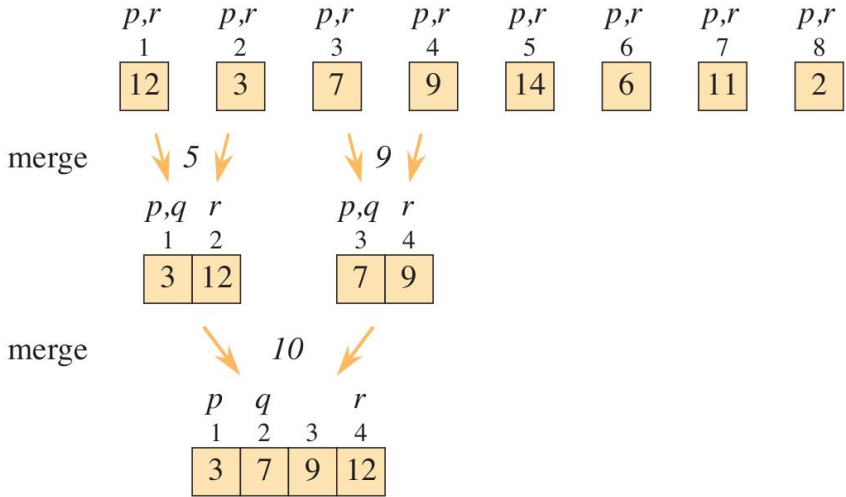
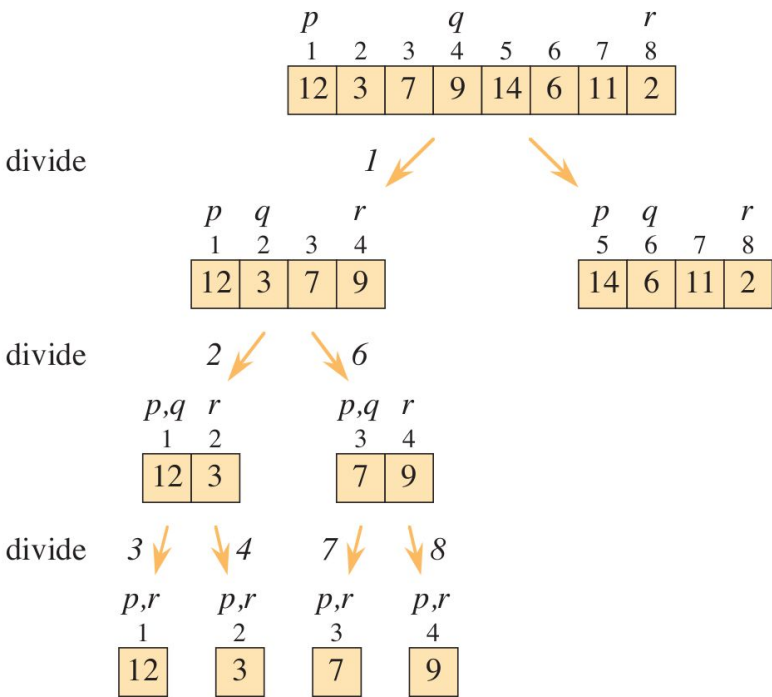
merge  $10 \searrow$



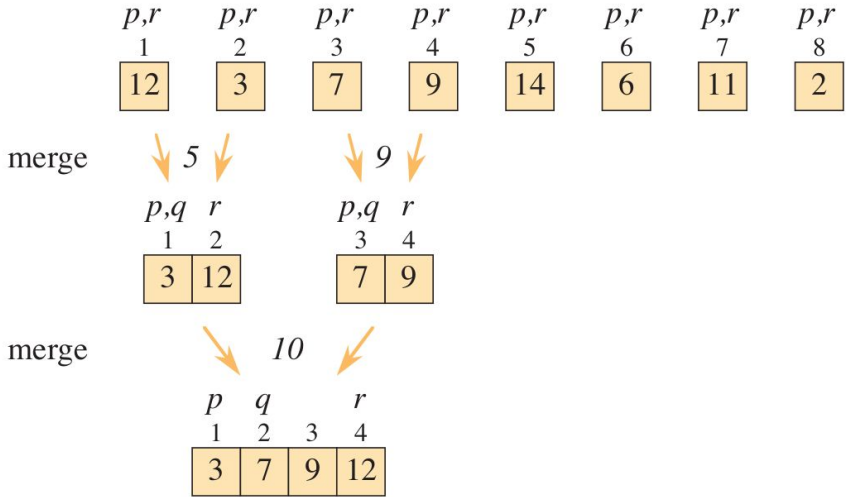
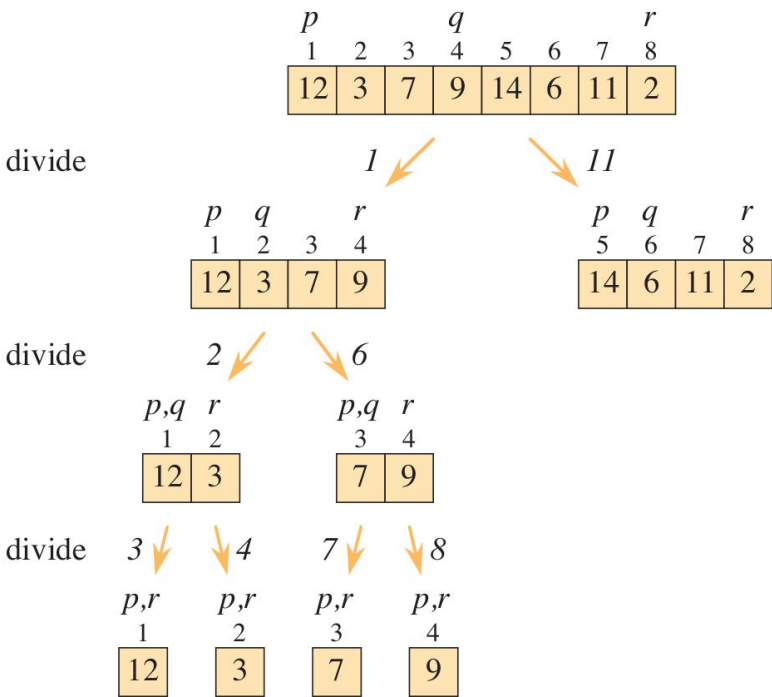
# MERGE SORT



# MERGE SORT

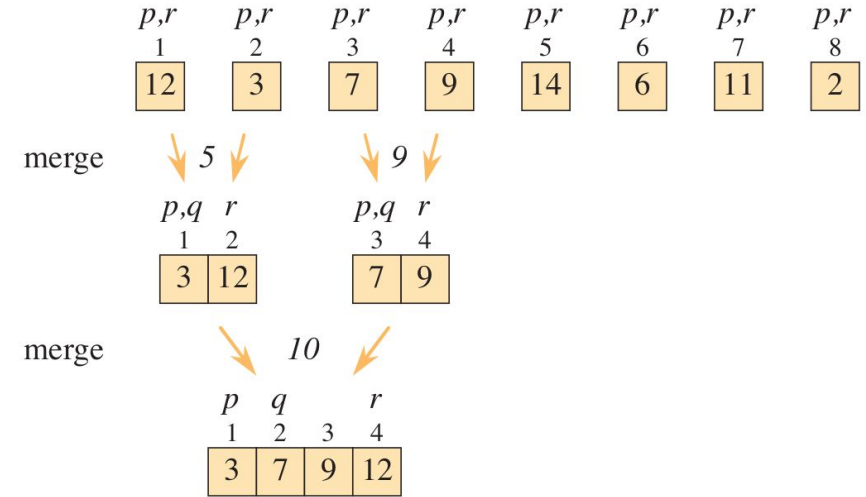
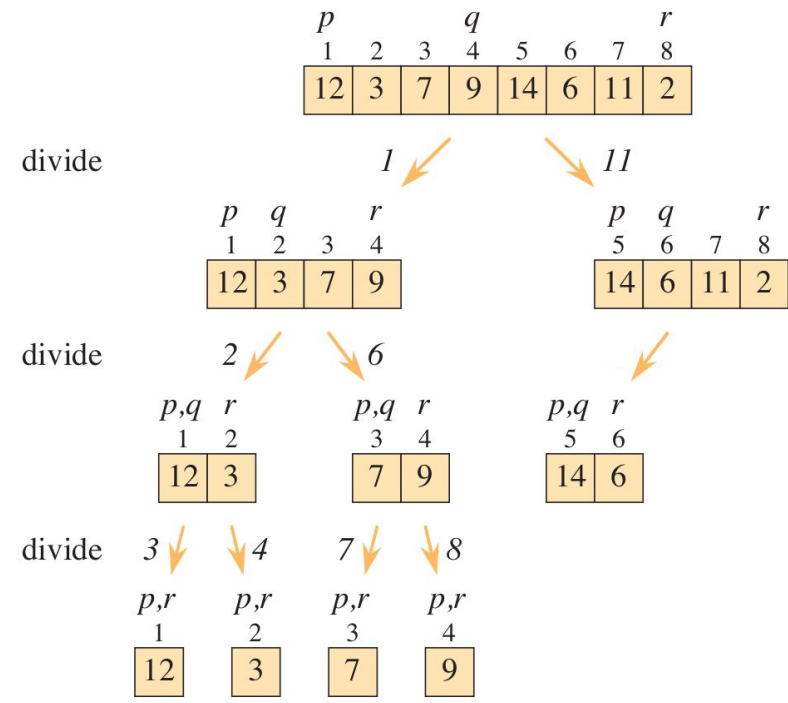


# MERGE SORT

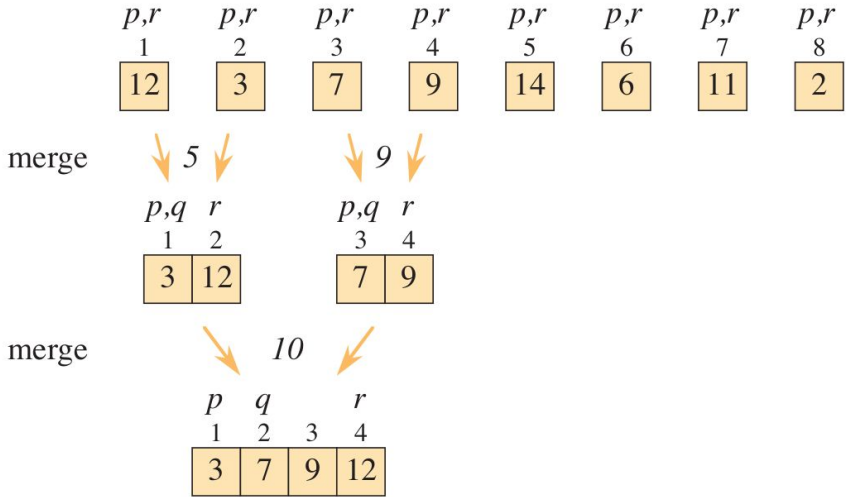
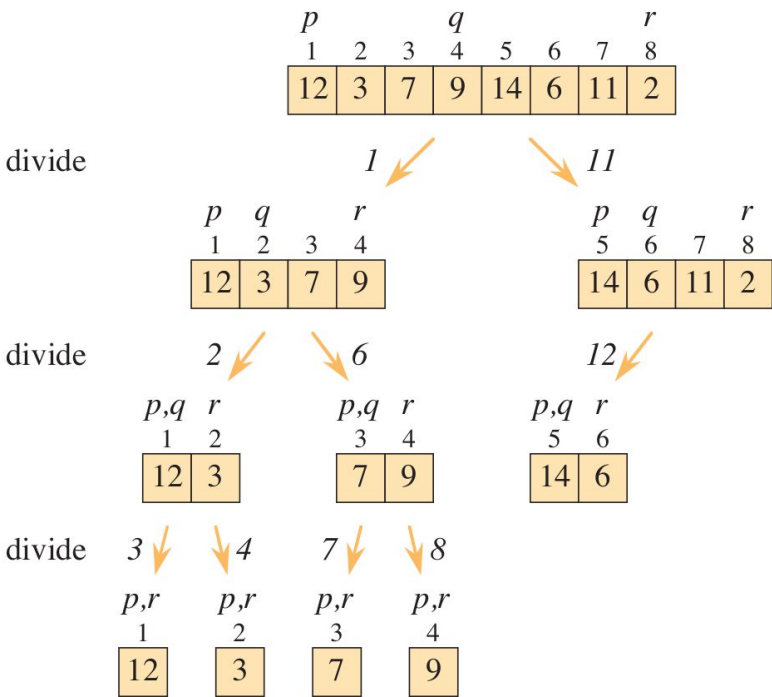




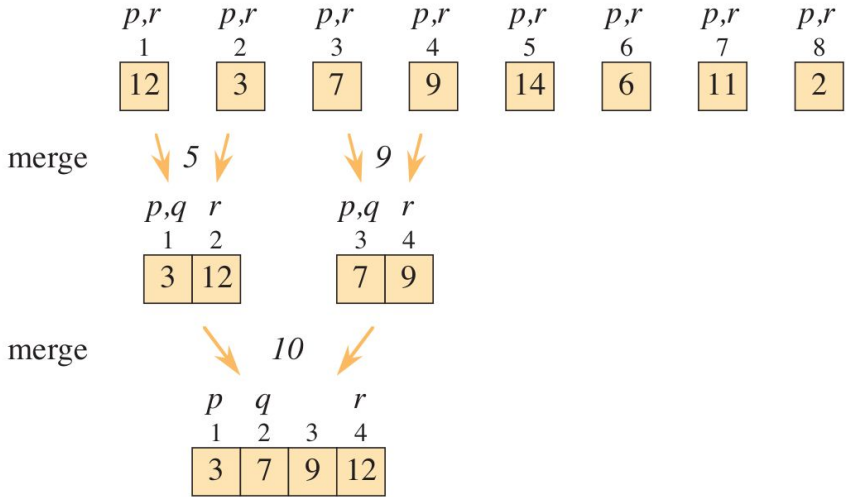
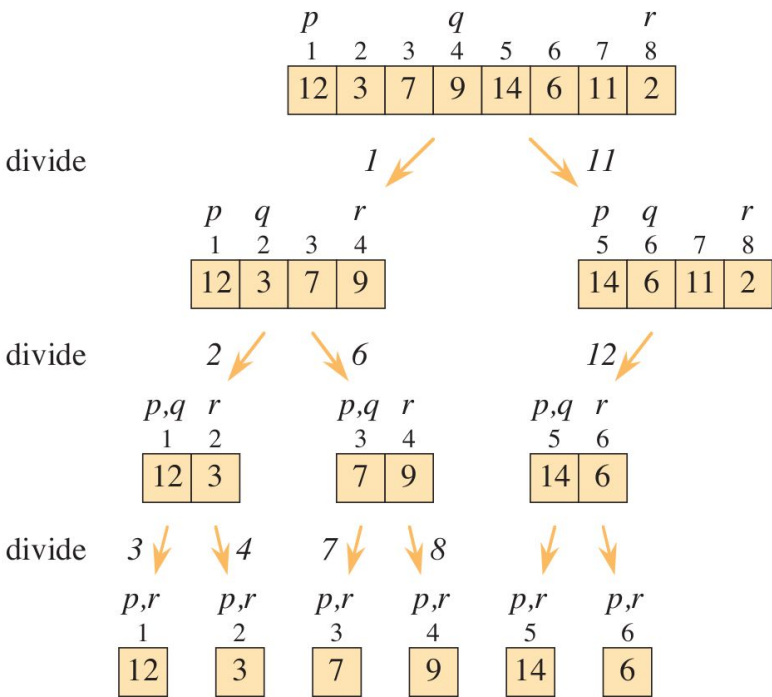
# MERGE SORT



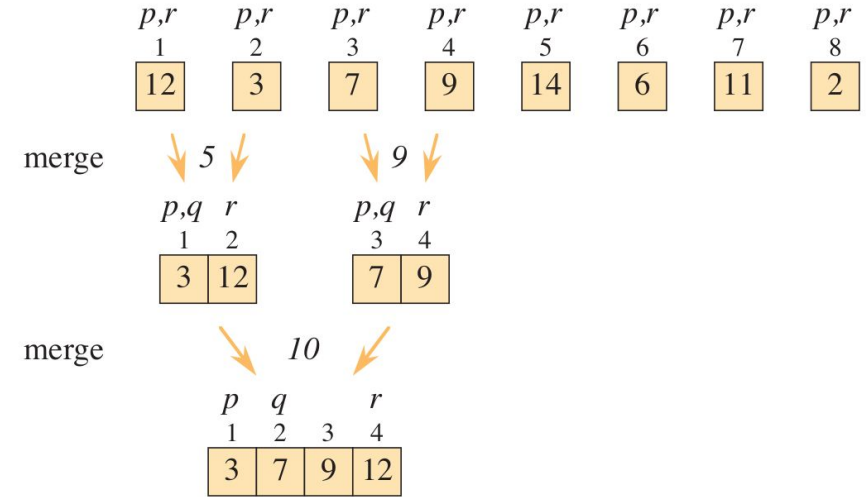
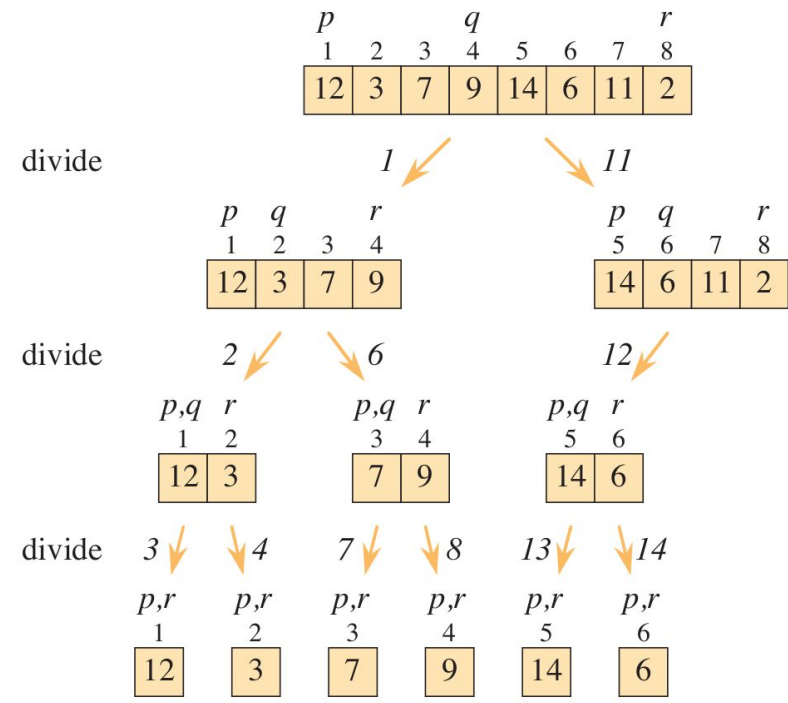
# MERGE SORT



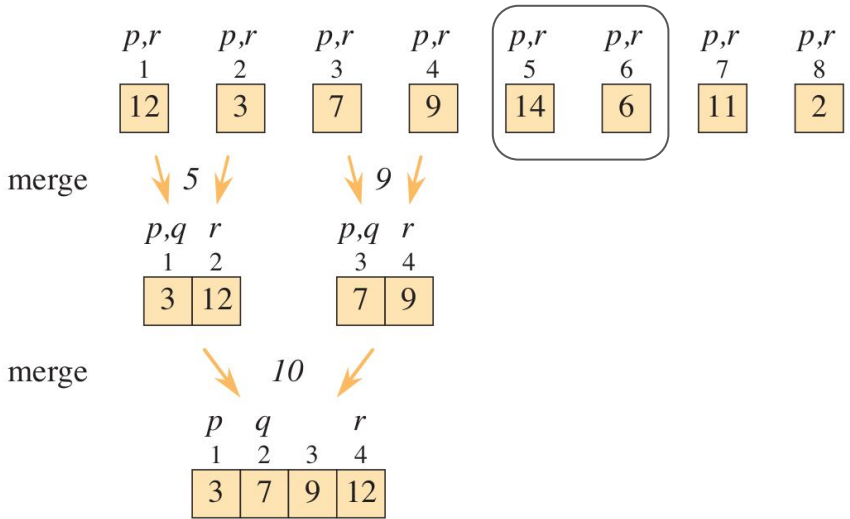
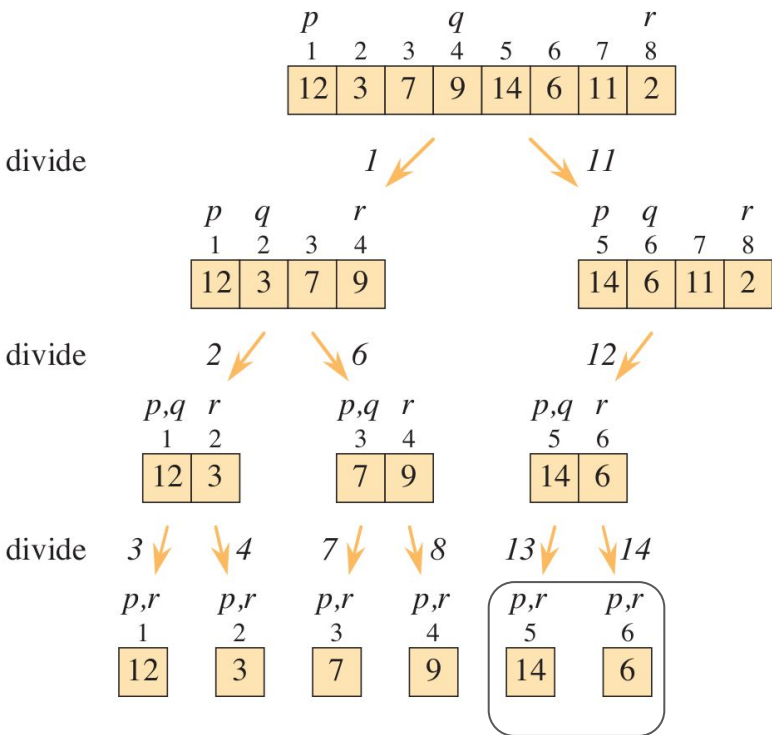
# MERGE SORT



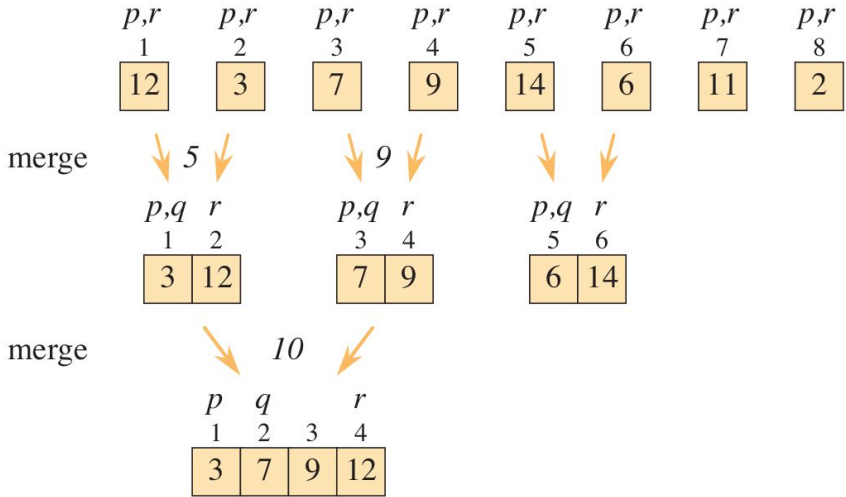
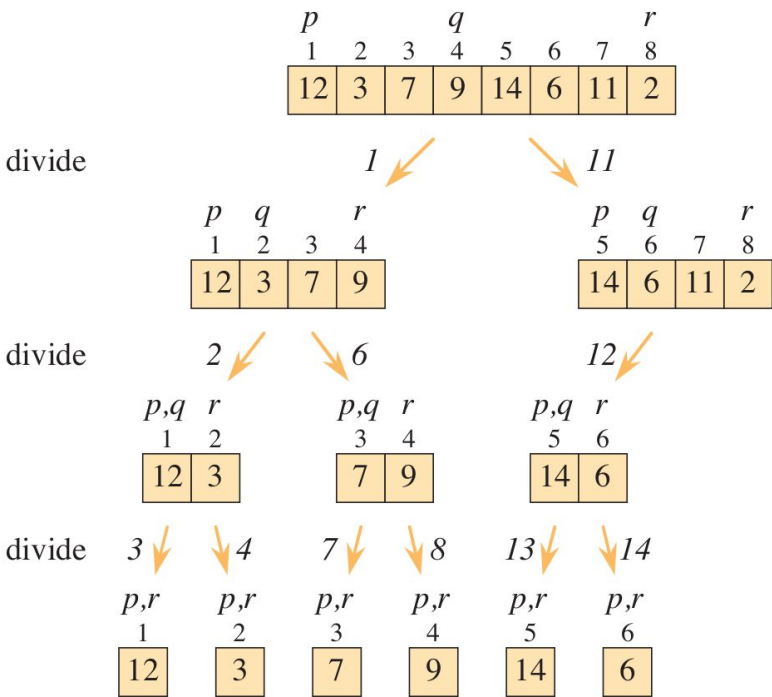
# MERGE SORT



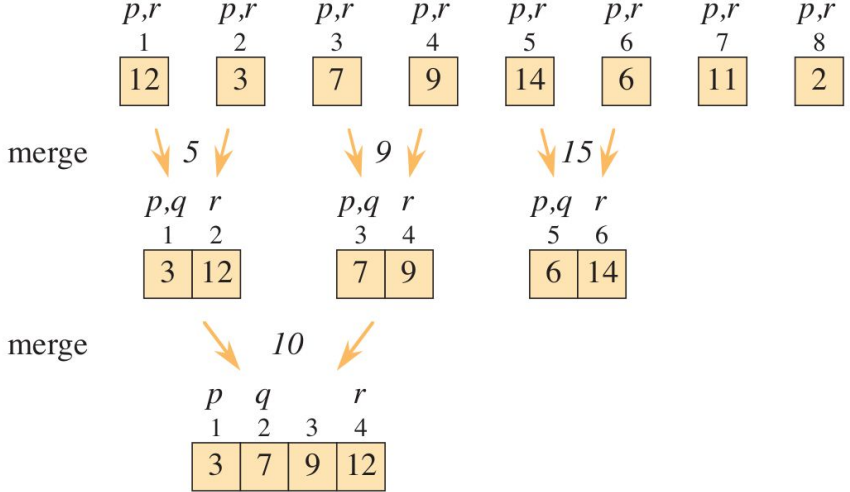
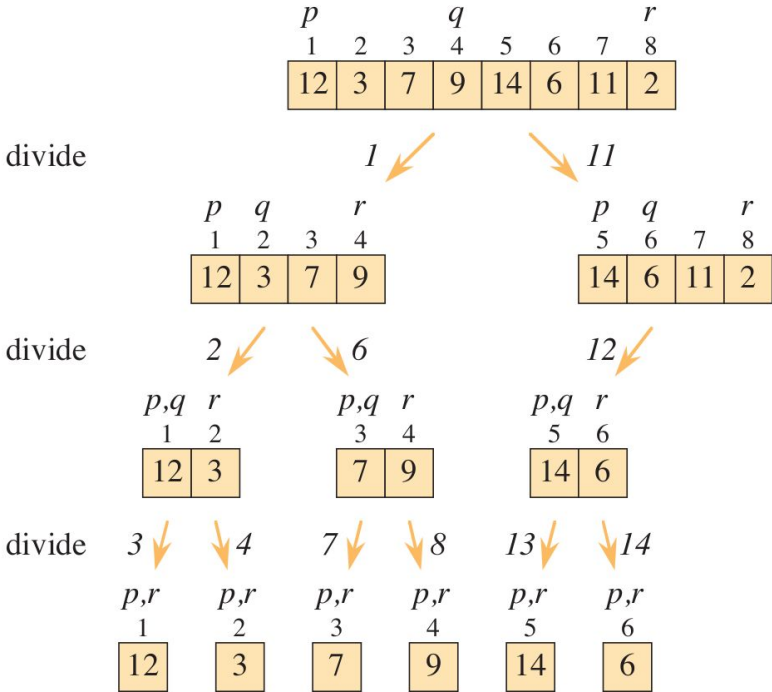
# MERGE SORT



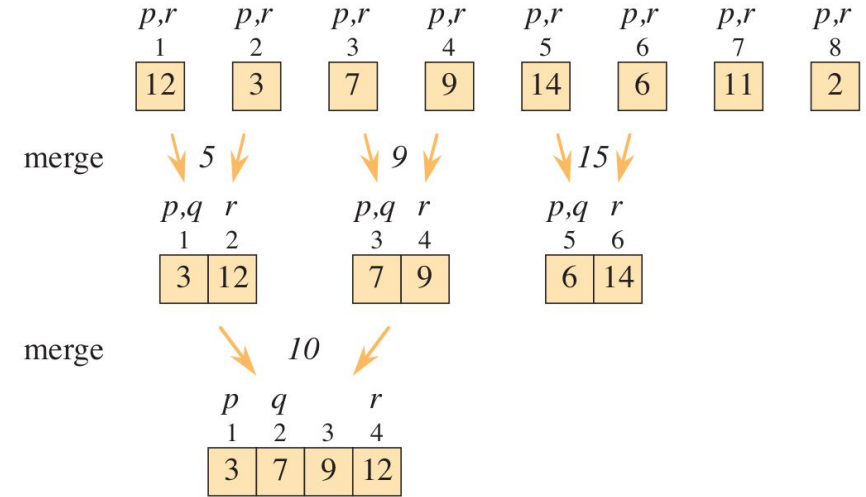
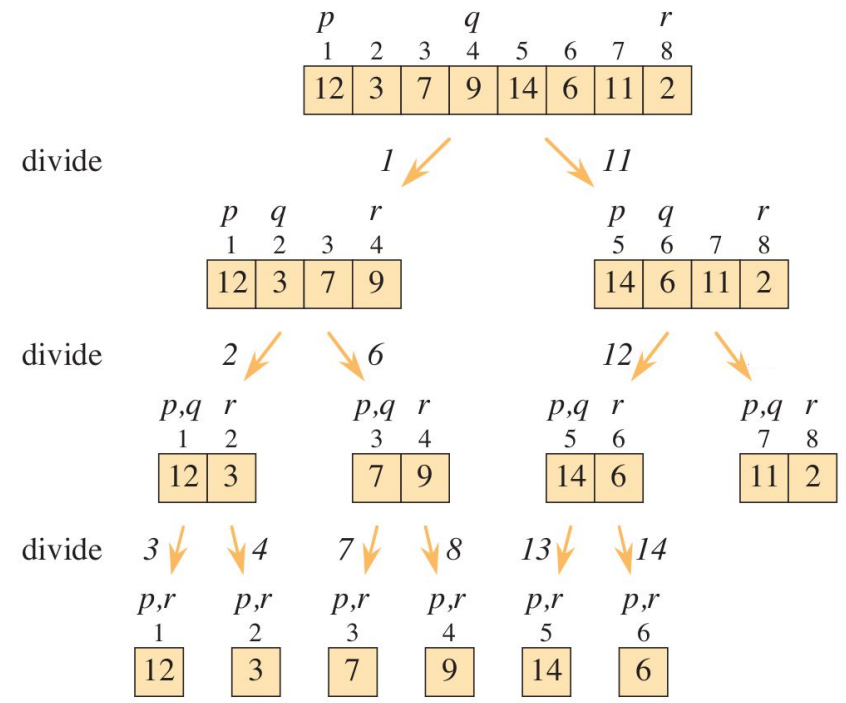
# MERGE SORT



# MERGE SORT

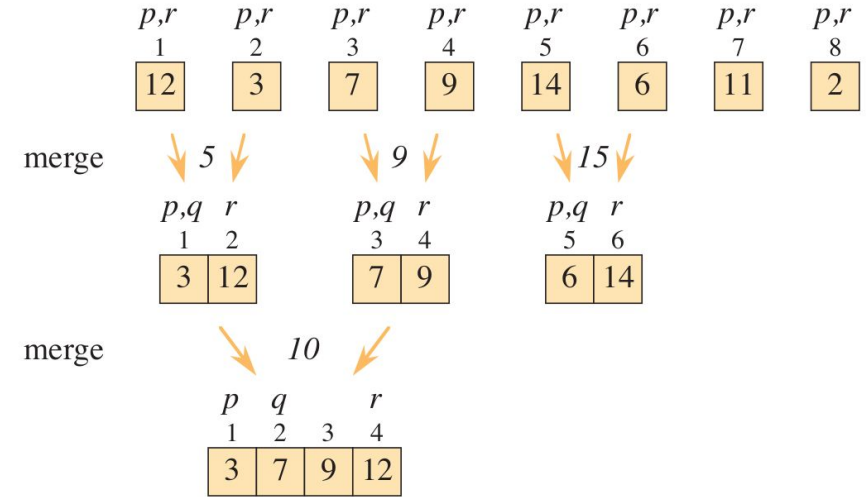
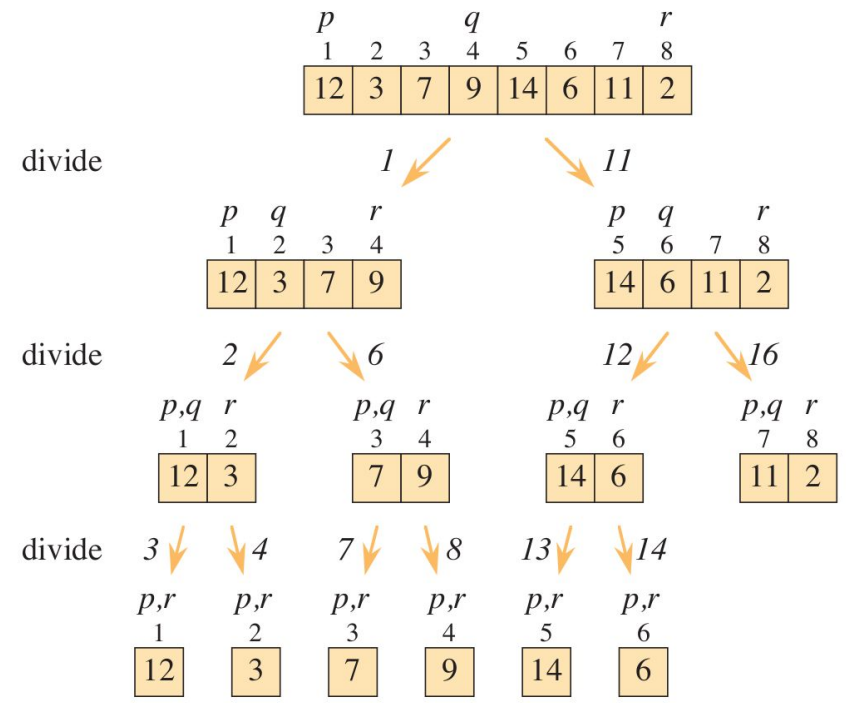


# MERGE SORT

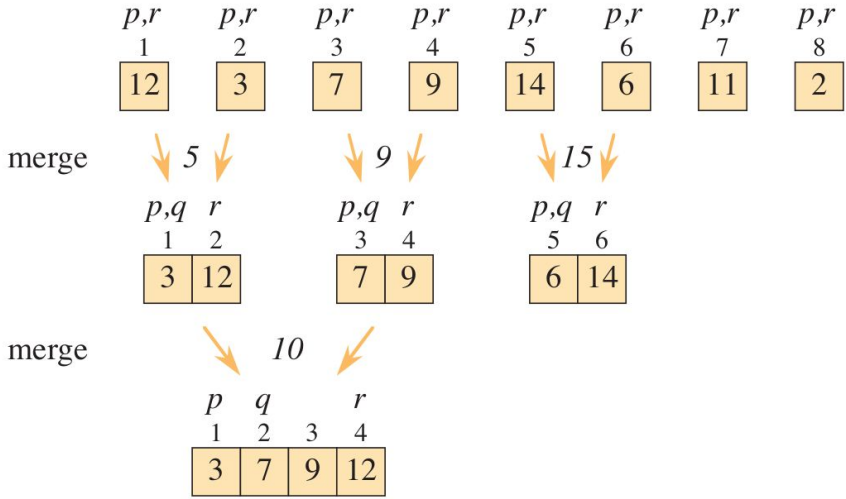
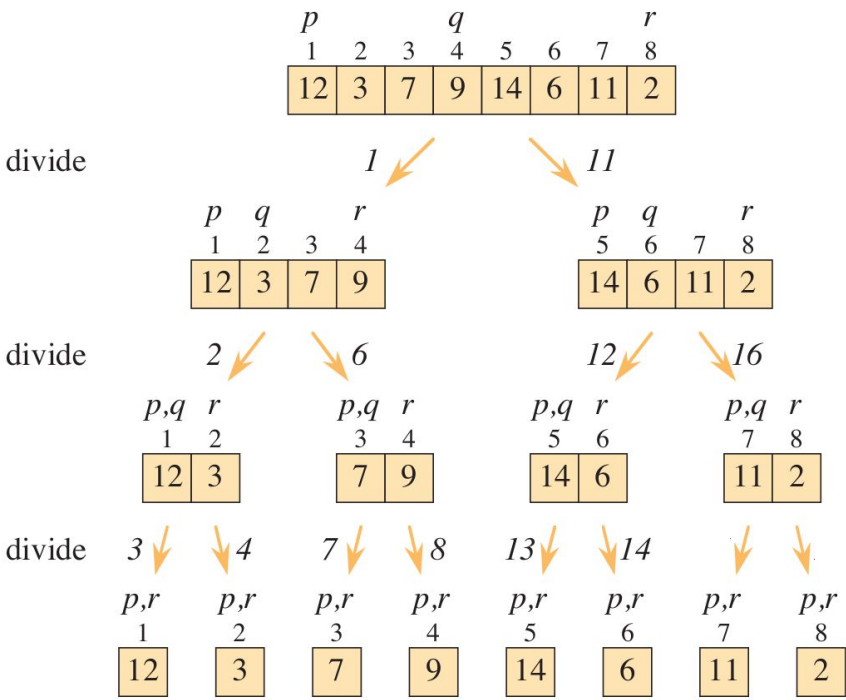




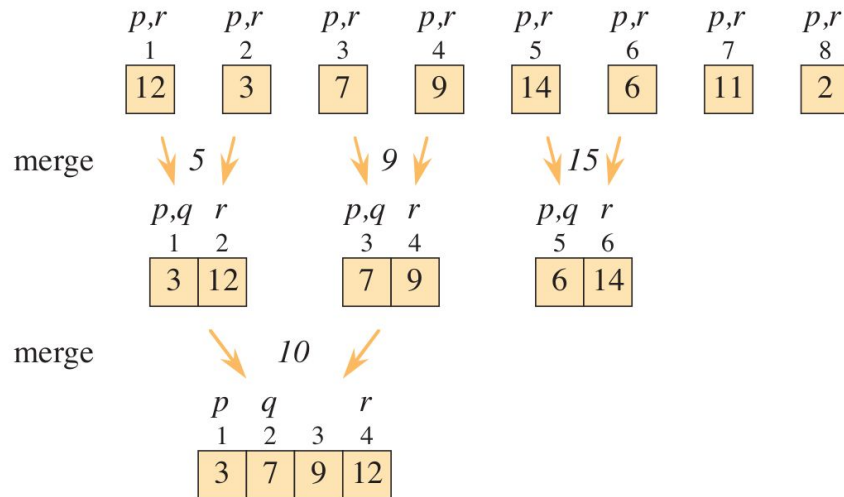
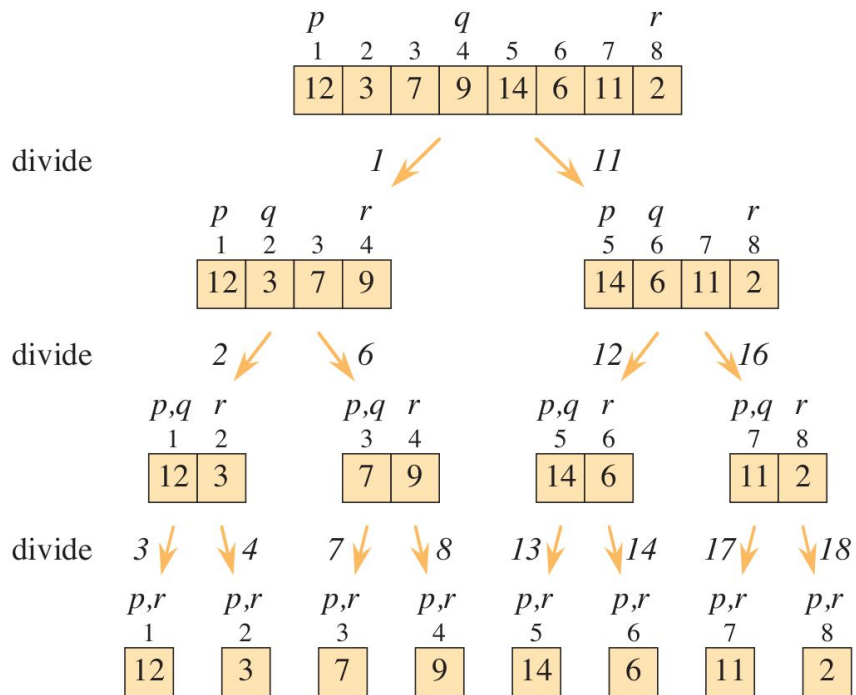
# MERGE SORT



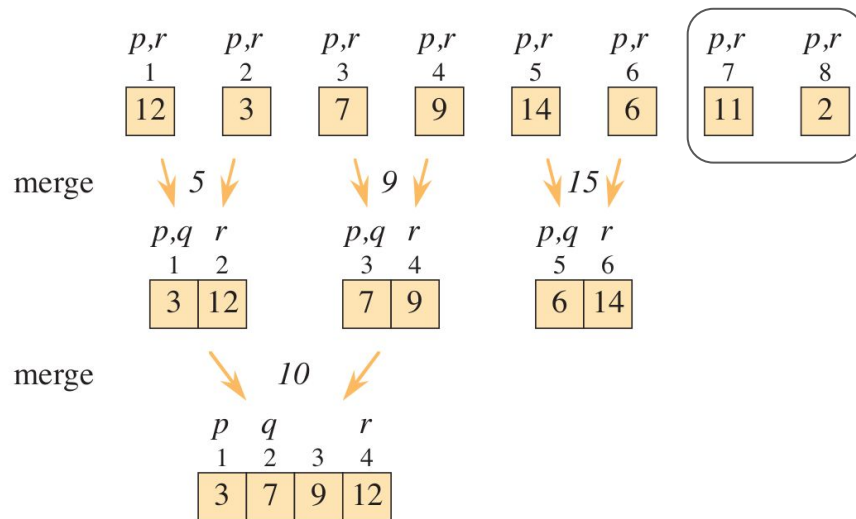
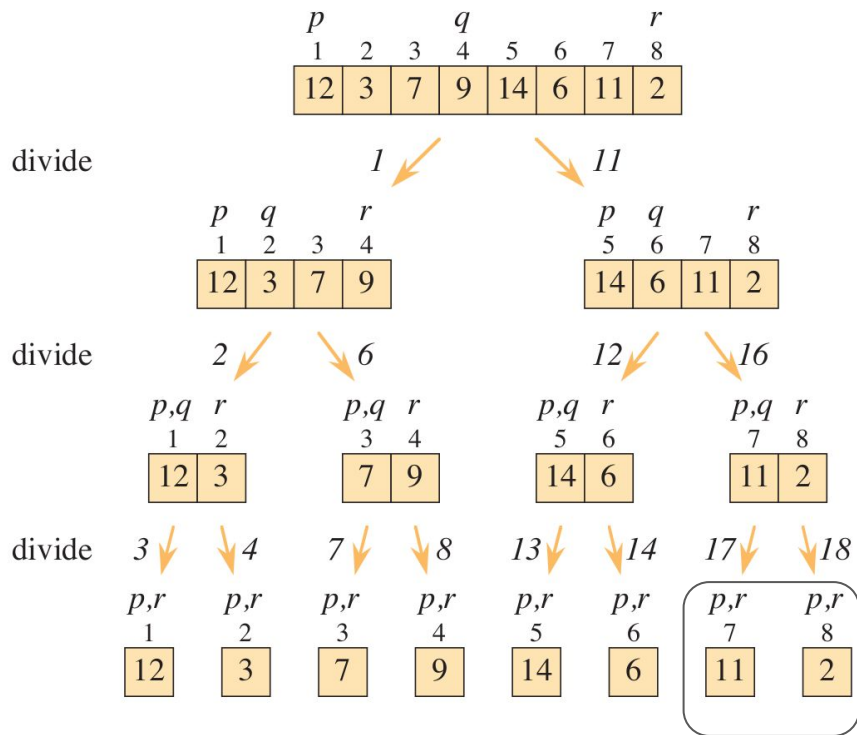
# MERGE SORT



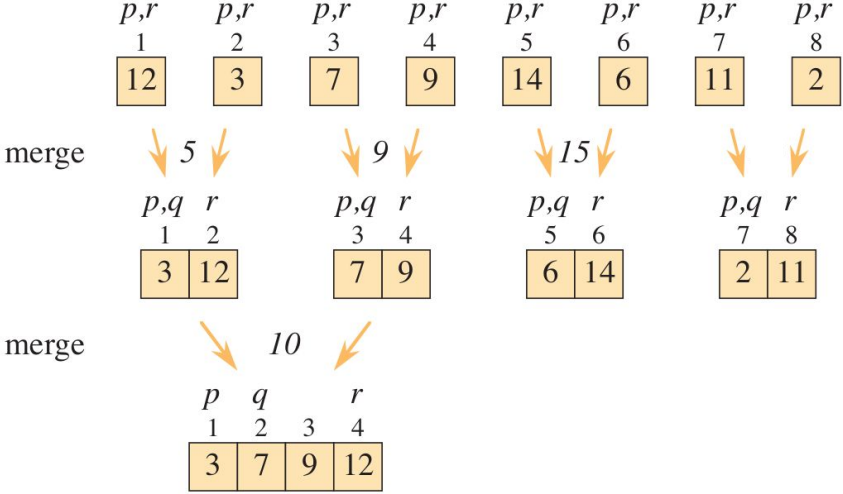
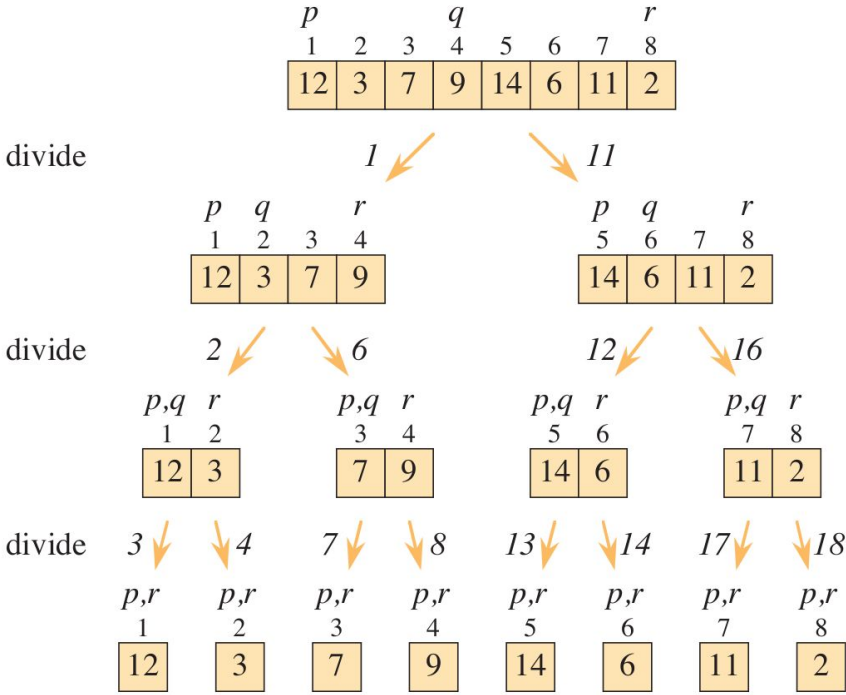
# MERGE SORT



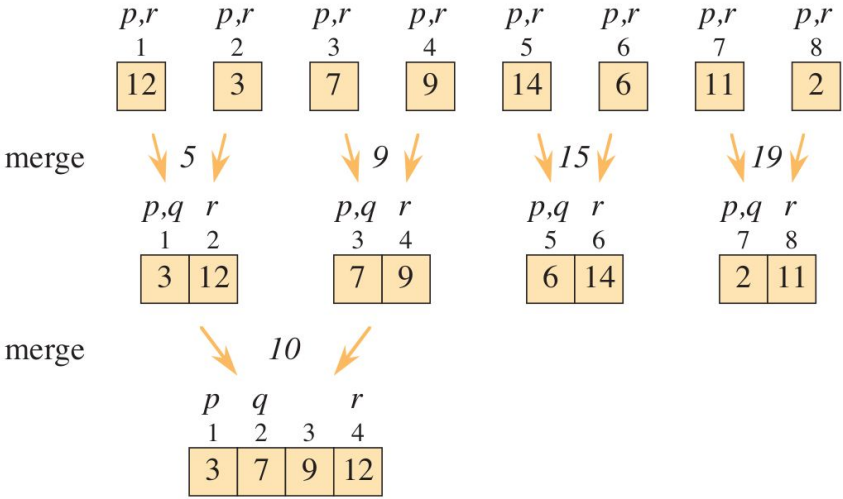
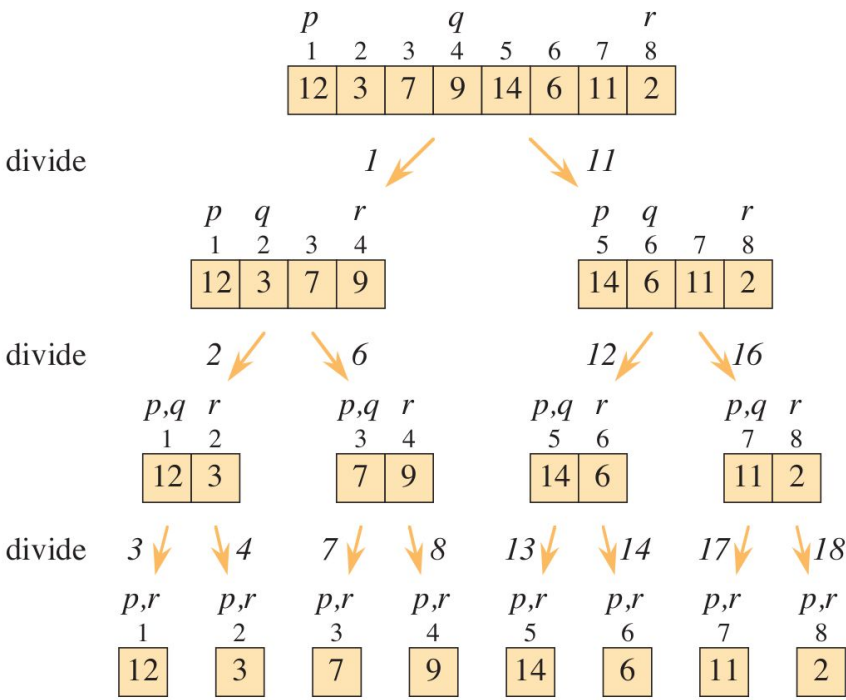
# MERGE SORT



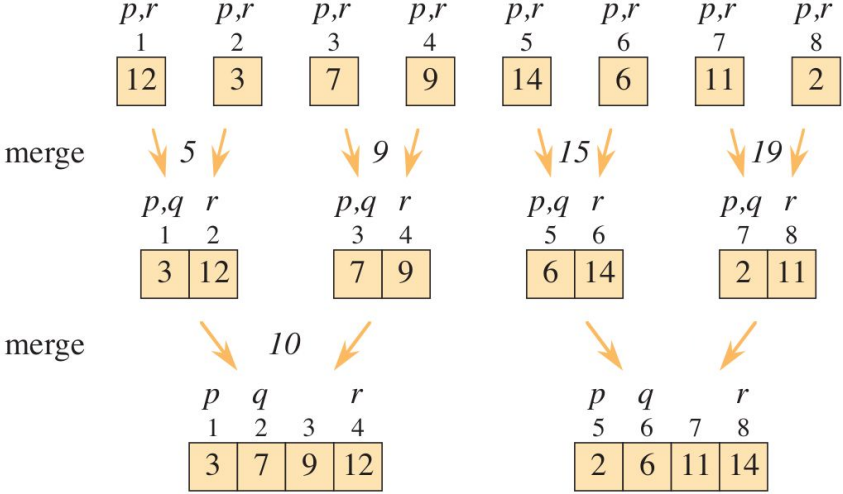
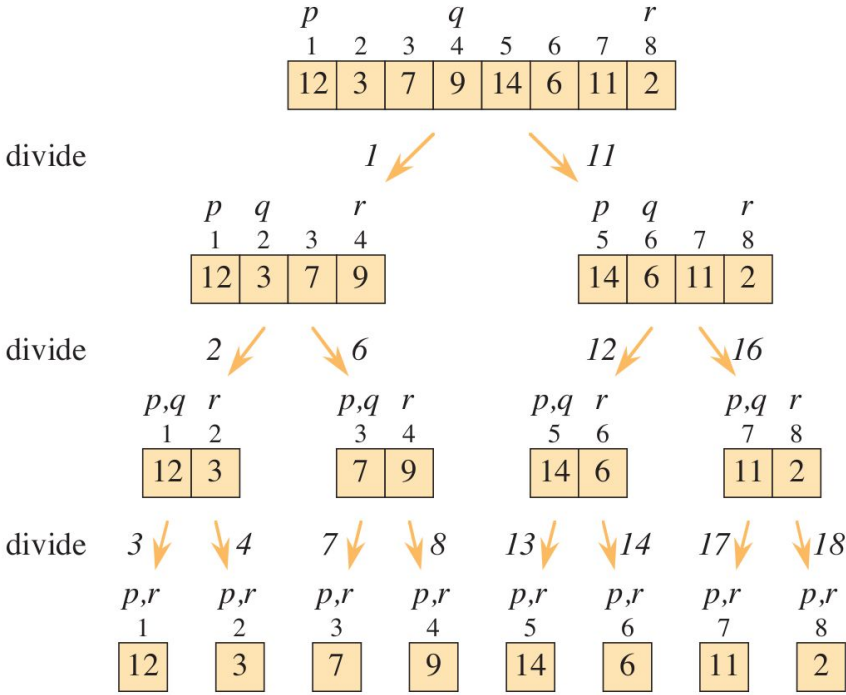
# MERGE SORT



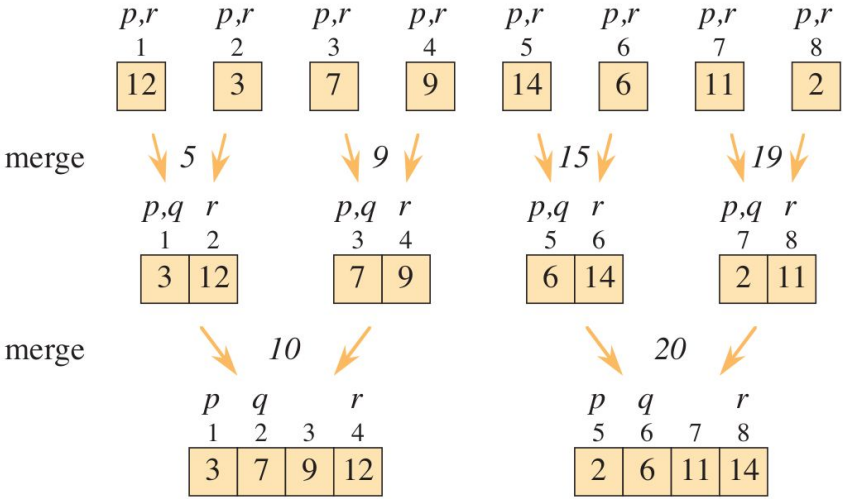
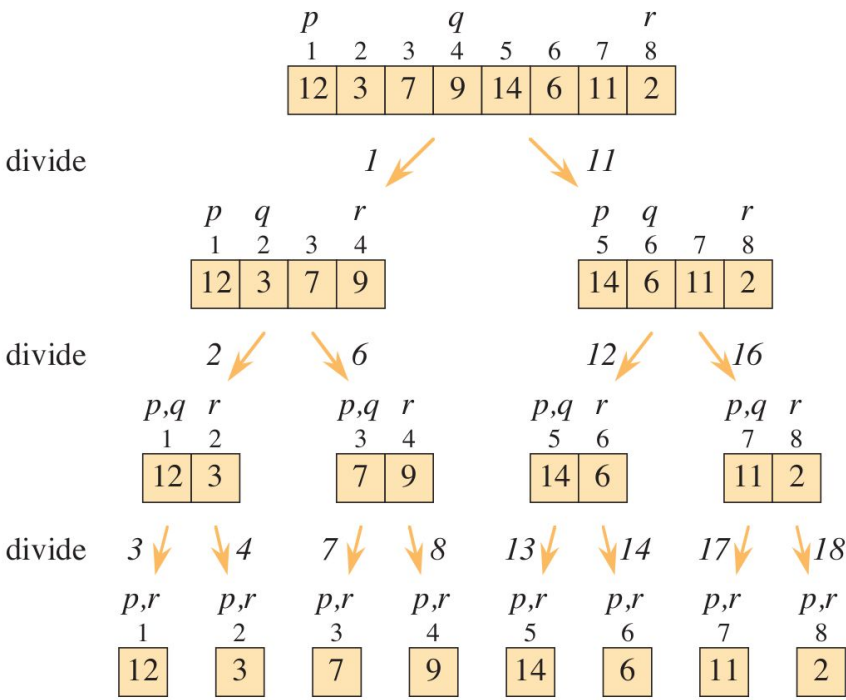
# MERGE SORT



# MERGE SORT

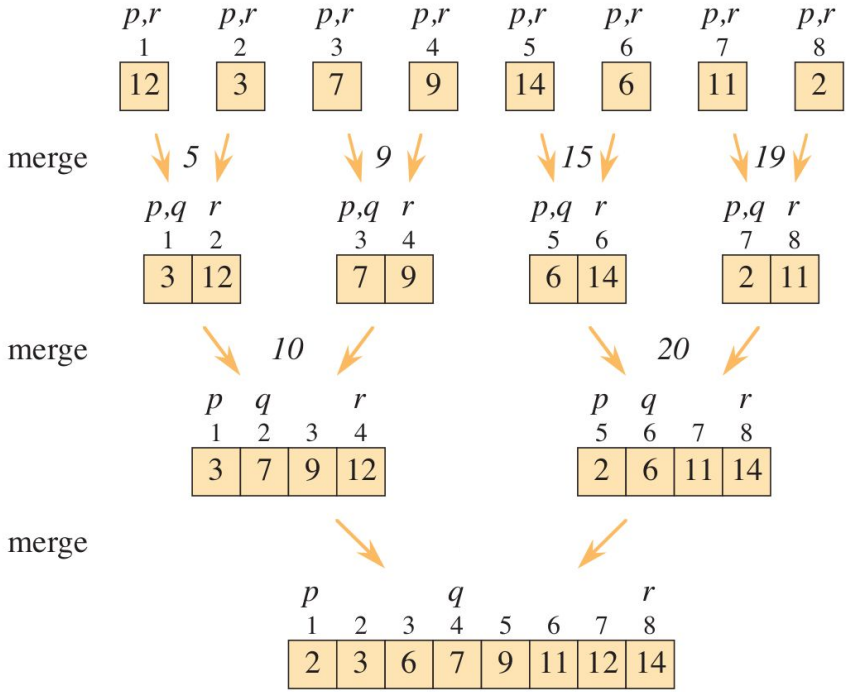
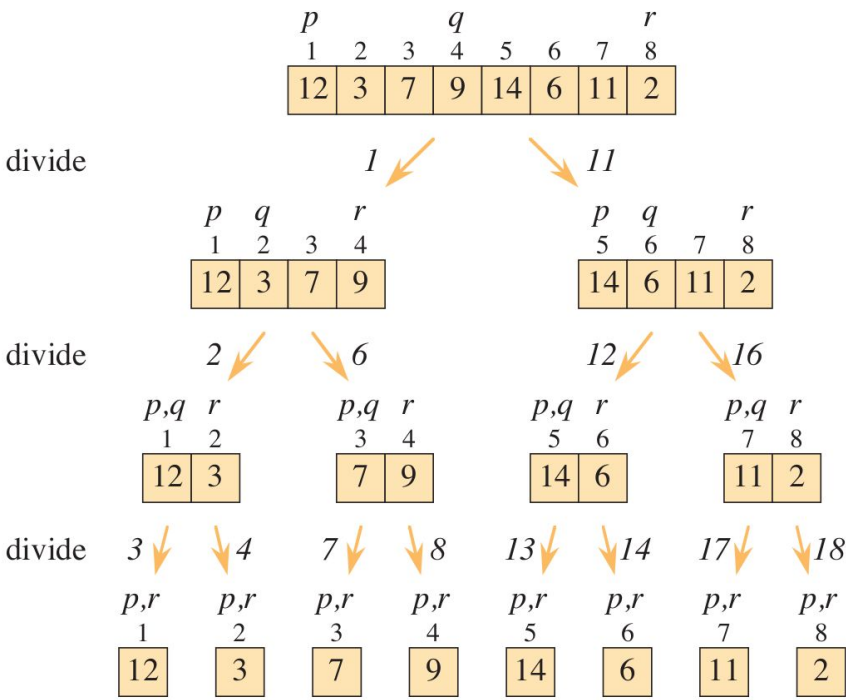


# MERGE SORT

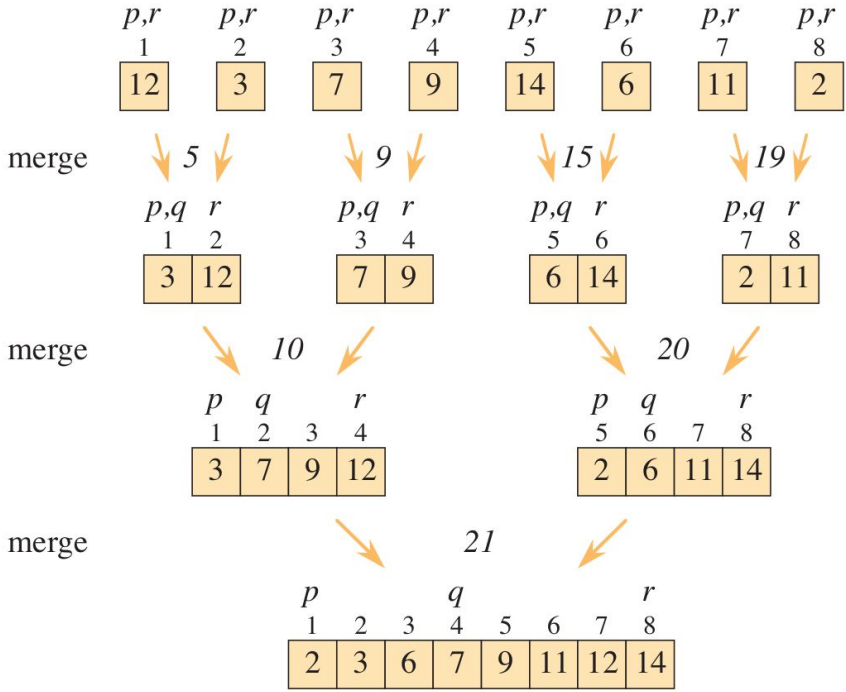
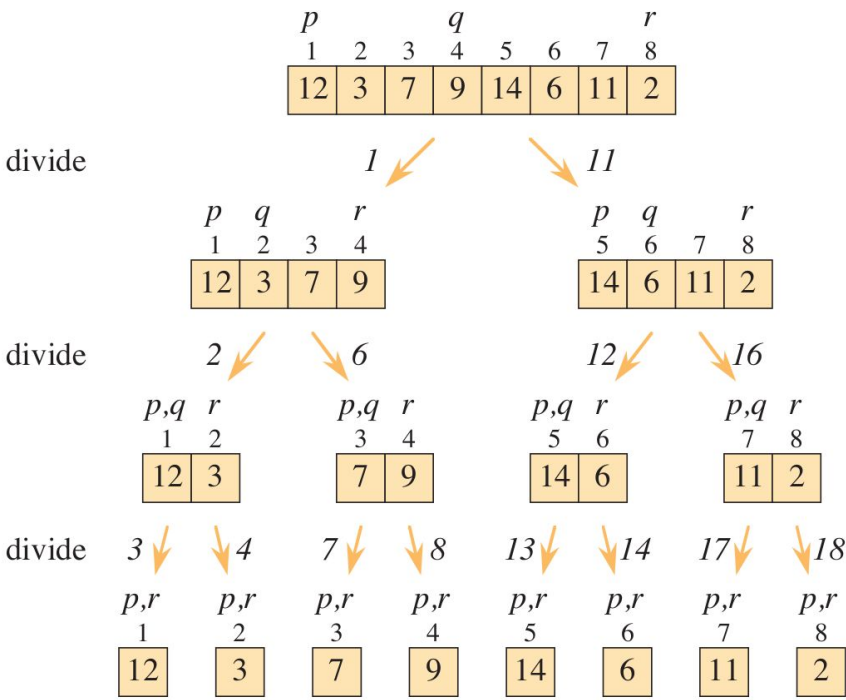




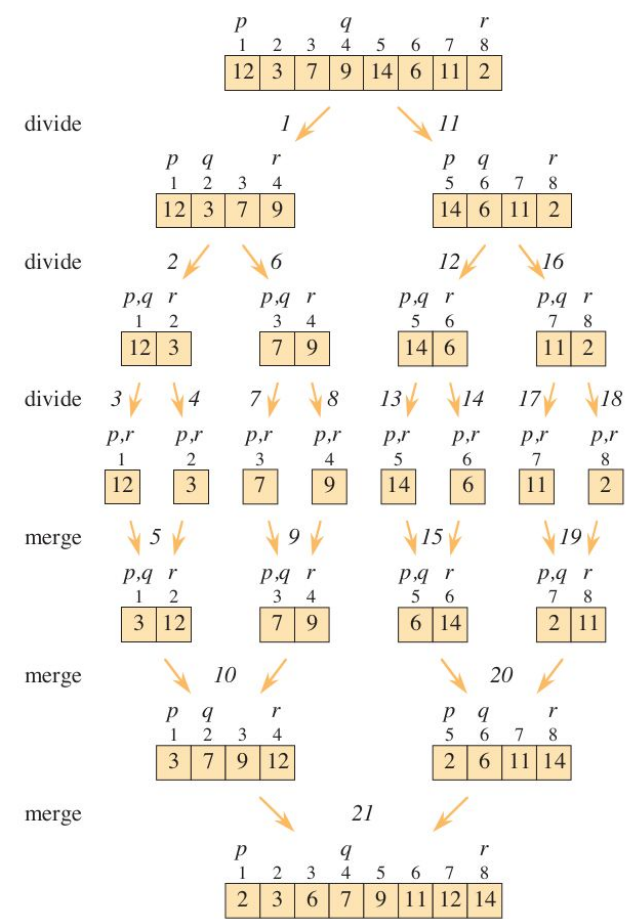
# MERGE SORT



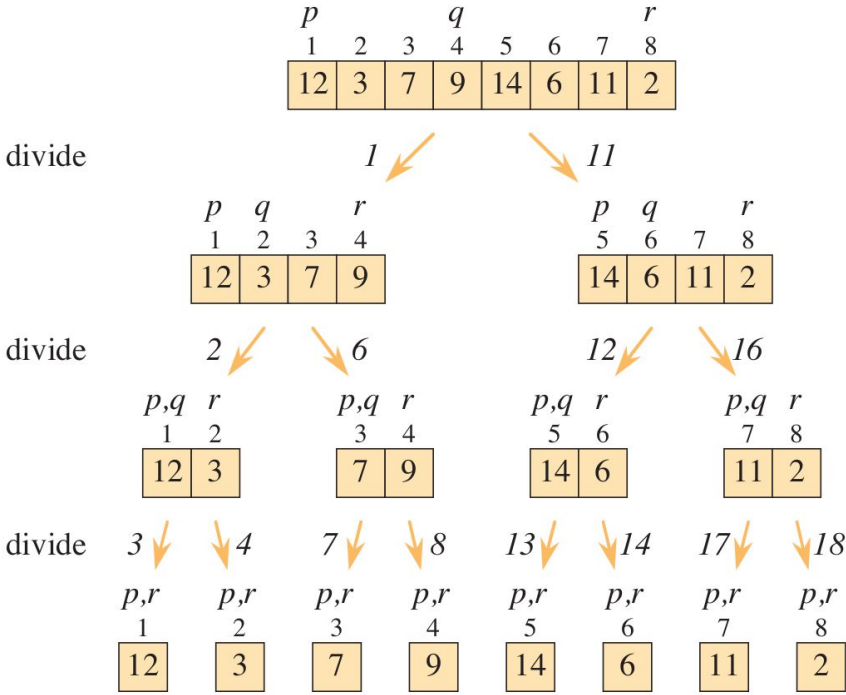
# MERGE SORT



# MERGE SORT



# MERGE SORT



MERGE-SORT( $A, p, r$ )

**if**  $p \geq r$

**return**

$q = \lfloor (p + r) / 2 \rfloor$

MERGE-SORT( $A, p, q$ )

MERGE-SORT( $A, q + 1, r$ )

// Merge  $A[p : q]$  and  $A[q + 1 : r]$  into  $A[p : r]$ .

MERGE( $A, p, q, r$ )

# MERGE SORT

**MERGE**( $A, p, q, r$ )

$n_L = q - p + 1$       // length of  $A[p : q]$

$n_R = r - q$       // length of  $A[q + 1 : r]$

let  $L[0 : n_L - 1]$  and  $R[0 : n_R - 1]$  be new arrays

**for**  $i = 0$  **to**  $n_L - 1$     // copy  $A[p : q]$  into  $L[0 : n_L - 1]$

$L[i] = A[p + i]$

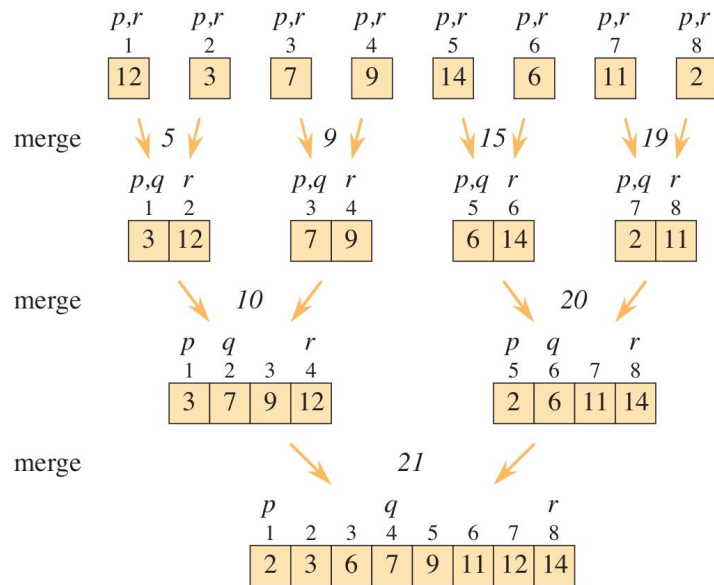
**for**  $j = 0$  **to**  $n_R - 1$     // copy  $A[q + 1 : r]$  into  $R[0 : n_R - 1]$

$R[j] = A[q + j + 1]$

$i = 0$       //  $i$  indexes the smallest remaining element in  $L$

$j = 0$       //  $j$  indexes the smallest remaining element in  $R$

$k = p$       //  $k$  indexes the location in  $A$  to fill



# MERGE SORT

// As long as each of the arrays  $L$  and  $R$  contains an unmerged element,

// copy the smallest unmerged element back into  $A[p:r]$ .

**while**  $i < n_L$  and  $j < n_R$

**if**  $L[i] \leq R[j]$

$A[k] = L[i]$

$i = i + 1$

**else**  $A[k] = R[j]$

$j = j + 1$

$k = k + 1$

// Having gone through one of  $L$  and  $R$  entirely, copy the

// remainder of the other to the end of  $A[p:r]$ .

**while**  $i < n_L$

$A[k] = L[i]$

$i = i + 1$

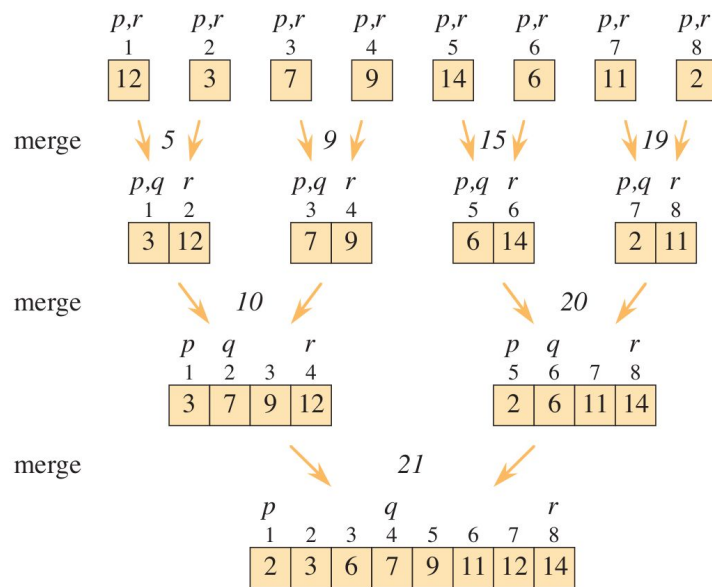
$k = k + 1$

**while**  $j < n_R$

$A[k] = R[j]$

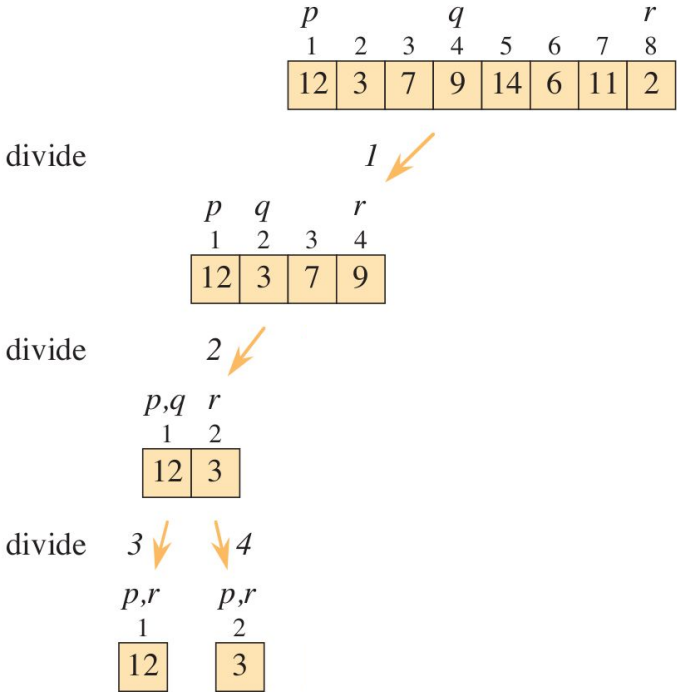
$j = j + 1$

$k = k + 1$



# MERGE SORT - ANALYSIS

# MERGE SORT



MERGE-SORT( $A, p, r$ )

```
if  $p \geq r$ 
    return
 $q = \lfloor (p + r) / 2 \rfloor$ 
MERGE-SORT( $A, p, q$ )
MERGE-SORT( $A, q + 1, r$ )
// Merge  $A[p : q]$  and  $A[q + 1 : r]$  into  $A[p : r]$ .
MERGE( $A, p, q, r$ )
```

$\log_2 8 = 3 \rightarrow 2^3 = 8$

$\log_2 4 = 2 \rightarrow 2^2 = 4$

$\log_b a = c \rightarrow b^c = a$



# MERGE SORT - ANALYSIS

$$T(n) = T(\frac{n}{2}) + T(\frac{n}{2}) + n$$

Splitting                      Merging

$$T(n) = 2T(\frac{n}{2}) + n$$

**Recursively apply the definition of  $T(n)$  to  $T(n/2)$**

$$T(\frac{n}{2}) = 2T(\frac{n}{4}) + \frac{n}{2}$$

# MERGE SORT - ANALYSIS

$$T(n) = T(\frac{n}{2}) + T(\frac{n}{2}) + n$$

Splitting

Merging

$$T(n) = 2T(\frac{n}{2}) + n$$

**Recursively apply the definition of  $T(n)$  to  $T(n/2)$**

$$T(\frac{n}{2}) = 2T(\frac{n}{4}) + \frac{n}{2}$$

$$T(n) = 2T(\frac{n}{2}) + n$$

$$T(n) = 2 \left( 2T(\frac{n}{4}) + \frac{n}{2} \right) + n$$

$$T(n) = \left( 2 \cdot 2T(\frac{n}{4}) + 2 \frac{n}{2} \right) + n$$

$$T(n) = 2^2 T(\frac{n}{4}) + n + n$$

$$T(n) = 2^2 T(\frac{n}{4}) + 2n$$

$$T(n) = 2^2 T(\frac{n}{2^2}) + 2n$$

# MERGE SORT - ANALYSIS

$$T(n) = T(\frac{n}{2}) + T(\frac{n}{2}) + n$$

Splitting      Merging

$$T(n) = 2T(\frac{n}{2}) + n$$

**Recursively apply the definition of  $T(n)$  to  $T(n/2)$**

$$T(\frac{n}{2}) = 2T(\frac{n}{4}) + \frac{n}{2}$$

$$T(n) = 2T(\frac{n}{2}) + n$$

$$T(n) = 2 \left( 2T(\frac{n}{4}) + \frac{n}{2} \right) + n$$

$$T(n) = \left( 2 \cdot 2T(\frac{n}{4}) + 2 \frac{n}{2} \right) + n$$

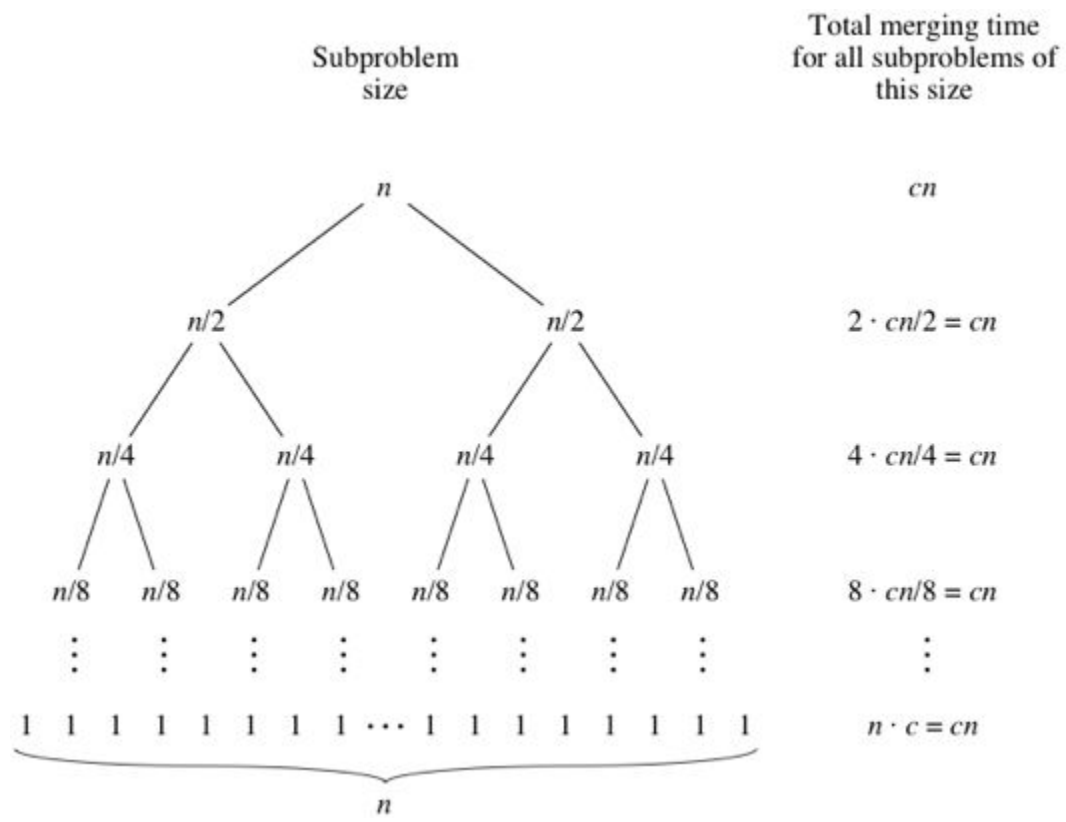
$$T(n) = 2^2 T(\frac{n}{4}) + n + n$$

$$T(n) = 2^2 T(\frac{n}{4}) + 2n$$

$$T(n) = 2^2 T(\frac{n}{2^2}) + 2n$$

$$T(n) = 2^i T(\frac{n}{2^i}) + i \cdot n$$

# MERGE SORT - ANALYSIS



# MERGE SORT - ANALYSIS

$$T(n) = 2^i T\left(\frac{n}{2^i}\right) + i \cdot n$$

$$T(1) = 1$$

$$T\left(\frac{n}{2^i}\right) = 1$$

$$\frac{n}{2^i} = 1$$

$$n = 2^i$$

$$\log_2 n = i$$

# MERGE SORT - ANALYSIS

$$T(n) = 2^i T\left(\frac{n}{2^i}\right) + i \cdot n$$

$$T(1) = 1$$

$$T\left(\frac{n}{2^i}\right) = 1$$

$$\frac{n}{2^i} = 1$$

$$n = 2^i$$

$$\log_2 n = i$$

$$T(n) = 2^i T\left(\frac{n}{2^i}\right) + i \cdot n$$

$$T(n) = nT(1) + \log_2 n \cdot n$$

$$T(n) = n + n \cdot \log_2 n$$

# MERGE SORT - ANALYSIS

$$T(n) = 2^i T\left(\frac{n}{2^i}\right) + i \cdot n$$

$$T(1) = 1$$

$$T\left(\frac{n}{2^i}\right) = 1$$

$$\frac{n}{2^i} = 1$$

$$n = 2^i$$

$$\log_2 n = i$$

$$T(n) = 2^i T\left(\frac{n}{2^i}\right) + i \cdot n$$

$$T(n) = nT(1) + \log_2 n \cdot n$$

$$T(n) = n + n \cdot \log_2 n$$

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

# INSERTION SORT



# ANALYZING ALGORITHMS

## *Best case*

The array is already sorted.

- Always find that  $A[i] \leq key$  upon the first time the **while** loop test is run (when  $i = i - 1$ ).
- All  $t_i$  are 1.
- Running time is

$$\begin{aligned} T(n) &= c_1n + c_2(n - 1) + c_4(n - 1) + c_5(n - 1) + c_8(n - 1) \\ &= (c_1 + c_2 + c_4 + c_5 + c_8)n - (c_2 + c_4 + c_5 + c_8) . \end{aligned}$$

- Can express  $T(n)$  as  $an + b$  for constants  $a$  and  $b$  (that depend on the statement costs  $c_k$ )  $\Rightarrow T(n)$  is a *linear function* of  $n$ .

# ANALYZING ALGORITHMS

**Worst case**

- Since  $\sum_{i=2}^n i = \left( \sum_{i=1}^n i \right) - 1$ , it equals  $\frac{n(n+1)}{2} - 1$ .
- Letting  $l = i - 1$ , we see that  $\sum_{i=2}^n (i - 1) = \sum_{l=1}^{n-1} l = \frac{n(n-1)}{2}$ .
- Running time is

$$\begin{aligned} T(n) &= c_1 n + c_2(n-1) + c_4(n-1) + c_5 \left( \frac{n(n+1)}{2} - 1 \right) \\ &\quad + c_6 \left( \frac{n(n-1)}{2} \right) + c_7 \left( \frac{n(n-1)}{2} \right) + c_8(n-1) \\ &= \left( \frac{c_5}{2} + \frac{c_6}{2} + \frac{c_7}{2} \right) n^2 + \left( c_1 + c_2 + c_4 + \frac{c_5}{2} - \frac{c_6}{2} - \frac{c_7}{2} + c_8 \right) n \\ &\quad - (c_2 + c_4 + c_5 + c_8) . \end{aligned}$$

- Can express  $T(n)$  as  $an^2 + bn + c$  for constants  $a, b, c$  (that again depend on statement costs)  $\Rightarrow T(n)$  is a *quadratic function* of  $n$ .

# COMPARISON

# MERGE SORT

## **Merge Sort**

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

## **Insertion Sort**

**Best:**  $an + b$

**Worst:**  $an^2 + bn + c$