F06 - Sortering 5DV149 Datastrukturer och algoritmer Kapitel 15

Niclas Börlin niclas.borlin@cs.umu.se

2020-02-06 Thu

Sortering

- Snabba upp andra algoritmer.
 - Sökning.
 - Hantera stora datamängder.



Sortering kontra Sorterad datatyp

- Sortering:
 - Förändrar ordningen mellan objekten i en struktur efter en sorteringsordning (fallande, ökande).
- Sorterad datatyp:
 - elementen i datatypen hålls sorterade enligt en sorteringsordning av de strukturförändrande operationerna i gränsytan (insert, delete).

Saker att beakta

- Absolut komplexitet:
 - Totala komplexiteten i alla implementationssteg
 - Kö (konstruerad som) lista (implementerad som) dubbellänkad lista.
- Passar en viss typ av sortering f\u00f6r en viss typ av implementation?
 - Fält-baserad lista
 - Länkad lista
- Ska man:
 - 1. sortera listan (tar tid) och sedan söka (går fortare)
 - 2. behålla listan osorterad (sparar tid) och göra linjär sökning (tar längre tid)?

Komplexitet

- ► Kan vara stor skillnad mellan värsta och bästa fallet.
- Fundera på:
 - ► Hur hanterar algoritmen en redan sorterad lista?
 - ► Hur hanterar algoritmen en motsatt sorterad lista?

Stabilitet

- Den inbördes relationen mellan två objekt med samma nyckel bibehålls vid sortering:
 - Lista sorterad efter första elementvärdet:

Lista omsorterad efter andra elementvärdet:

► Alla sorteringsalgoritmer går inte att göra stabila.

Grundprinciper

- Instickssortering:
 - Välj ett godtyckligt element och sätt in det på rätt plats. Ex: Insertion sort.
- Urvalssortering:
 - Välj ut det objekt som är på tur att sättas in. Sätt in det sist/först. Ex: Selection sort.
- Utbytessortering:
 - Byt plats på objekt som ligger fel inbördes. Ex: Bubble sort.
- Samsortering:
 - Sammanslagning av redan sorterade strukturer. Ex. Merge sort, Inplace Quicksort.
- (Nyckelsortering:)
 - Kräver mer information/kunskap om objektmängden.

Utbytessortering

- ► Utbytessorteringsalgoritmer (*inplace*-) kräver endast *O*(1) extra minne.
- List-algoritmer kan kräva O(n) extra minne.
- ► Flera algoritmer finns både som list-versioner och utbytesversioner.

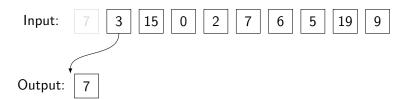
Sorteringsalgoritmer

- ► Idag:
 - ► Instickssortering (*Insertion Sort*)
 - Urvalssortering (Selection Sort)
 - Bubbelsortering (Bubble Sort)
 - Mergesort
 - Quicksort
- ► Senare:
 - Heapsort
 - (Radix Exchange Sort)

Input: 7 3 15 0 2 7 6 5 19 9



Input: 7 3 15 0 2 7 6 5 19 9



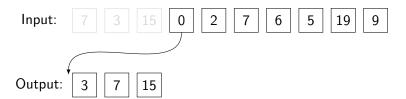
Input: 7 3 15 0 2 7 6 5 19 9

Output: 3 7



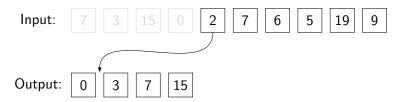
Input: 7 3 15 0 2 7 6 5 19 9

Output: 3 7 15



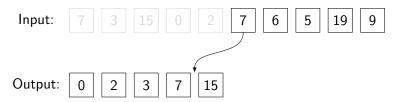
Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 3 7 15



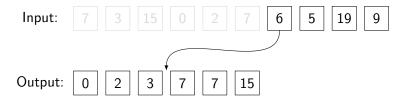
Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3 7 15



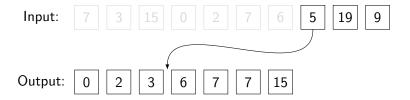
Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3 7 7 15



Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3 6 7 7 15



Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3 5 6 7 7 15





Output: 0 2 3 5 6 7 7 15 19



Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3 5 6 7 7 9 15 19

Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3 5 6 7 7 9 15 19

Input: 7 3 15 0 2 7 6 5 19 9

Input: 7 3 15 0 2 7 6 5 19 9





Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2



Output: 0 2

Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3

Input: 7 3 15 0 2 7 6 5 19 9

Output: 0 2 3



Output: 0 2 3 5



Output: 0 2 3 5



Output: 0 2 3 5 6



Output: 0 2 3 5 6



Output: 0 2 3 5 6 7



Output: 0 2 3 5 6 7



Output: 0 2 3 5 6 7 7



Output: 0 2 3 5 6 7 7



Output: 0 2 3 5 6 7 7 9



Output: 0 2 3 5 6 7 7 9



Output: 0 2 3 5 6 7 7 9 15



Output: 0 2 3 5 6 7 7 9 15



Output: 0 2 3 5 6 7 7 9 15 19

Bubbelsortering

```
Algorithm Bubblesort(arr)
repeat
    swapped \leftarrow false
    for j \leftarrow low(arr) to high(arr)-1 do
        if arr[j] > arr[j+1] then
            temp \leftarrow arr[j]
            arr[j] \leftarrow arr[j+1]
            arr[j+1] ← temp
            swapped <- true
until not swapped
 Stabil sortering?
 Tidskomplexiteten?
      \triangleright O(n^2) för en fältbaserad lista.
      ► O(?) för en länkad lista?
```

Divide-and-Conquer

- Rekursiv algoritmprincip:
 - Grundidé: Dela upp problemet i mindre och mindre problem tills problemet är trivialt.
 - Lös det triviala basfallet.
 - Slå ihop till en totallösning.
- ► Mergesort och Quicksort är av denna typ.
- ► Komplexitet: $O(n \log n)$.

Mergesort

- ► Algoritm för att sortera sekvensen S:
- ▶ Om S har bara ett element (trivialt):
 - ► Sekvensen S redan sorterad. Returnera S.

annars

- ▶ Divide: Dela S i två lika stora delsekvenser S1 och S2.
 - ► Sortera sekvenserna S1 och S2 rekursivt.
- Conquer: Slå samman S1 och S2 till en sorterad sekvens S. Returnera S.

Algoritm Mergesort (1)

```
Algorithm Mergesort(S)

if length(S) > 1 then
   (S1,S2) \( \times \) Split(S)
   S1 \( \times \) Mergesort(S1)
   S2 \( \times \) Mergesort(S2)
   S \( \times \) Merge(S1,S2)

return S
```

Algoritm Mergesort (2)

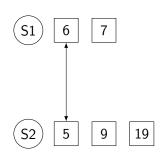
```
Algorithm Merge (S1,S2)
S ← Empty()
while not Isempty(S1) and not Isempty(S2) do
   if Inspect(first(S1),S1) \le Inspect(first(S2),S2) then
      Insert(Inspect(first(S1),S1),end(S),S)
      Remove(First(S1),S1)
   else
      Insert(Inspect(first(S2),S2),end(S),S)
      Remove(First(S2),S2)
while not Isempty(S1) do
   Insert(Inspect(First(S1),S1),end(S),S)
   Remove(First(S1),S1)
while not Isempty(S2) do
   Insert(Inspect(First(S2),S2),end(S),S)
   Remove(First(S2),S2)
return S
```

S1 6 7

(s)

S2) 5 9 19





S1 6 7

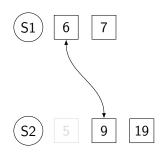
(S) 5

S2

5

19



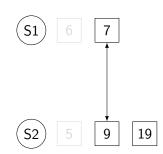


S1 6 7

 $\left(\mathsf{S}\right)\left[\mathsf{S}\right]\left[\mathsf{6}\right]$

S2 5 9 19





S1 6 7

S 5 6 7

S2) 5 9 19

S1 6 7

S2 5 9 19

S1 6 7

S 5 6 7 9 19

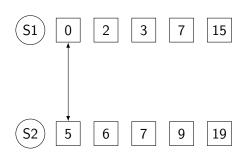
S2 5 9 19

 S1
 0
 2
 3
 7
 15

3)

S2 5 6 7 9 19

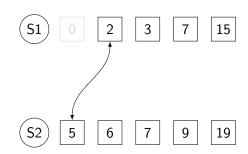




S1 0 2 3 7 15

 (S2)
 5
 6
 7
 9
 19

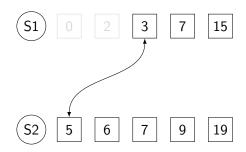




S1 0 2 3 7 15

S2 5 6 7 9 19



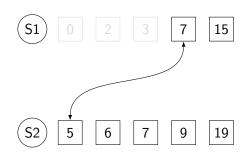


S1 0 2 3 7 15

 $\left(\mathsf{S} \right) \boxed{0} \boxed{2} \boxed{3}$

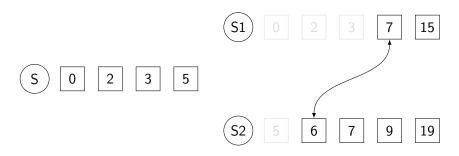
S2 5 6 7 9 19





S1 0 2 3 7 15

S2 5 6 7 9 19





(S) 0 2 3 5 6

S2) 5 6 7 9 19





S 0 2 3 5 6 7

S2 5 6 7 9 19





S 0 2 3 5 6 7 7

S2 5 6 7 9 19





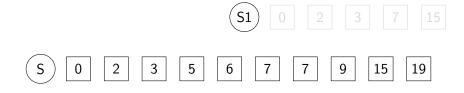
S 0 2 3 5 6 7 7 9

S2 5 6 7 9 19





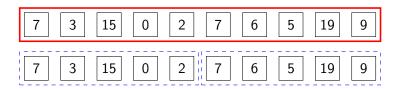
- (S) 0 2 3 5 6 7 7 9 15
 - S2 5 6 7 9 19

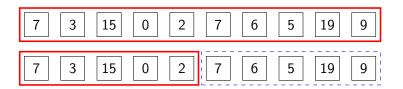


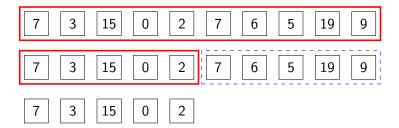
7 3 15 0 2 7 6 5 19 9

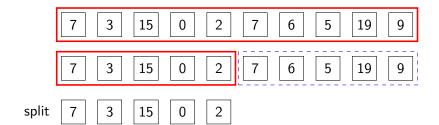


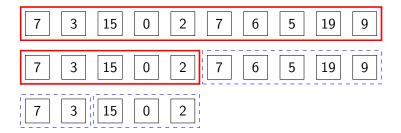


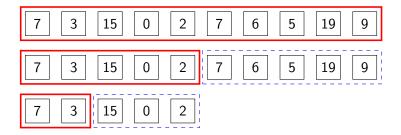


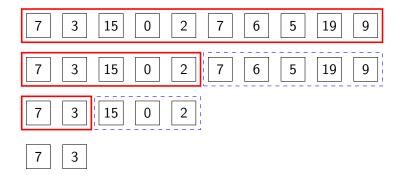


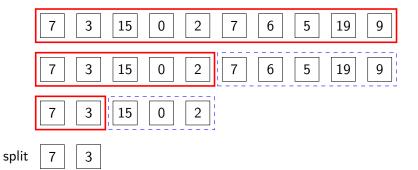


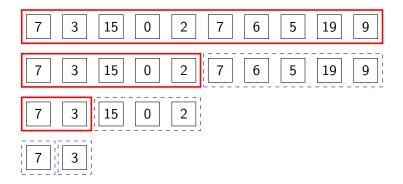


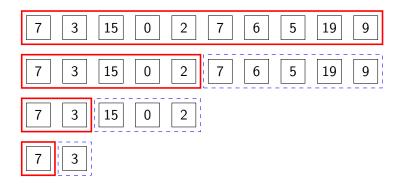


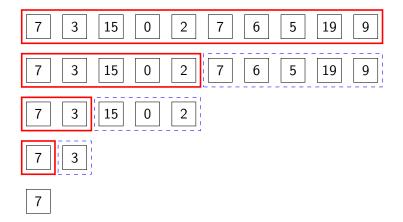


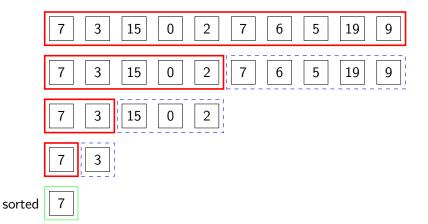


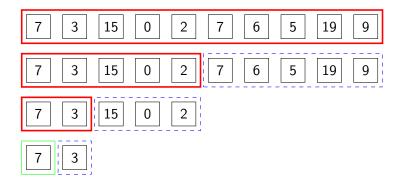


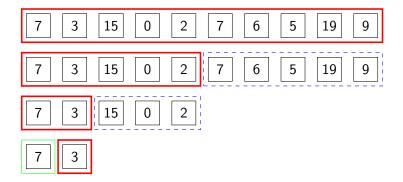


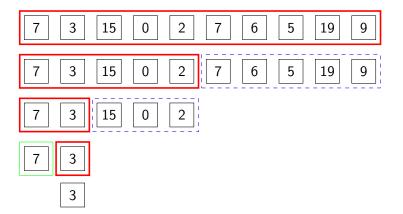


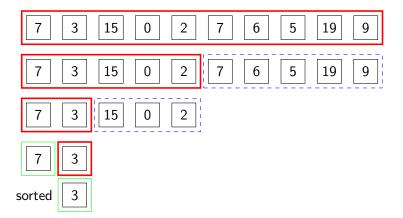


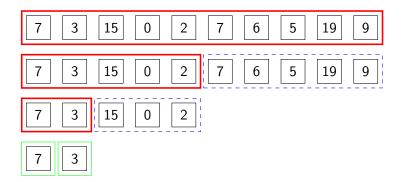


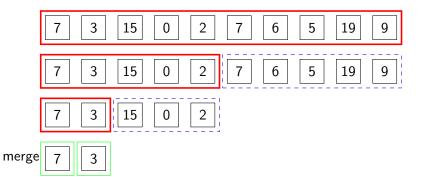


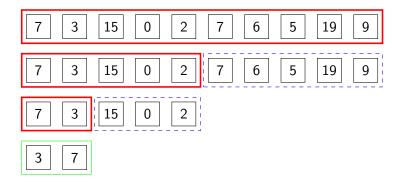


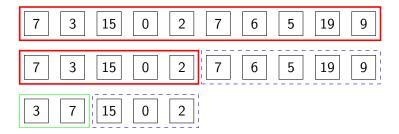


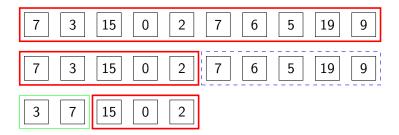


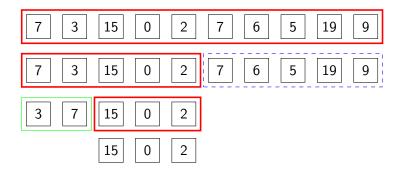


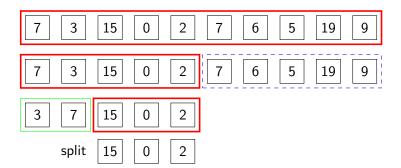


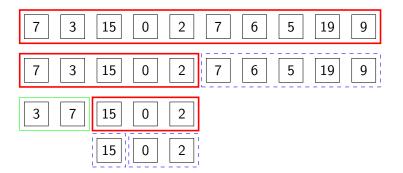


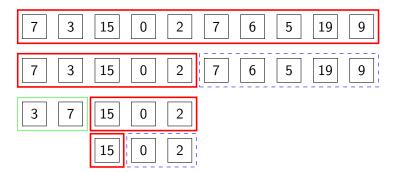


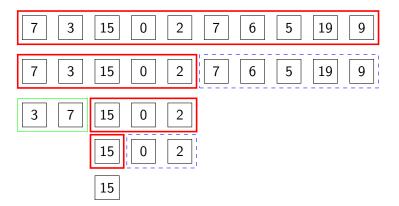


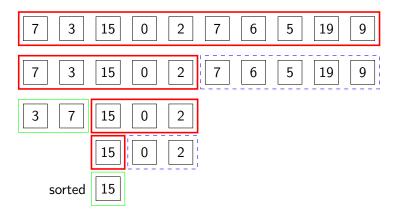


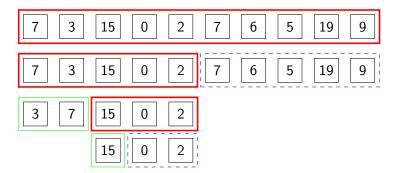


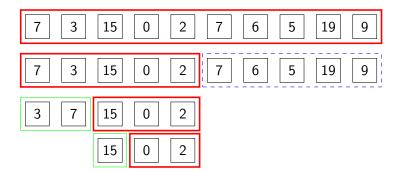


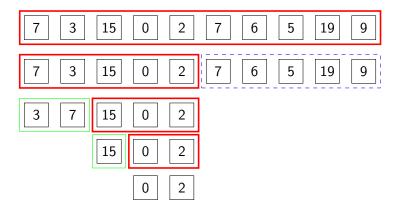


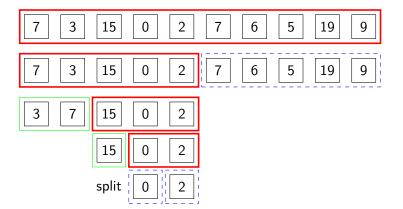


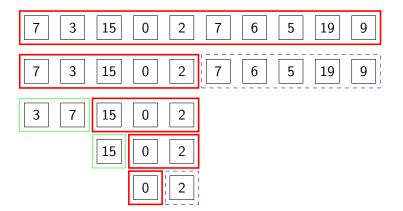


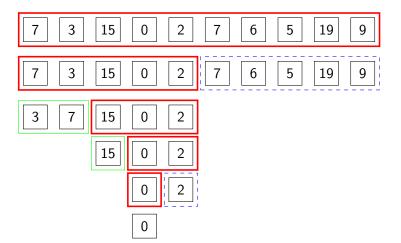


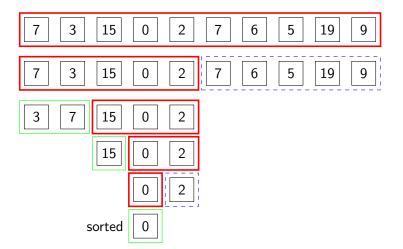


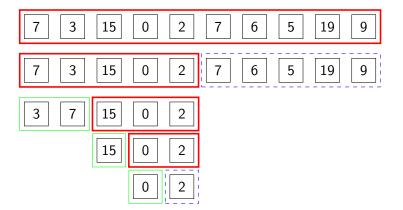


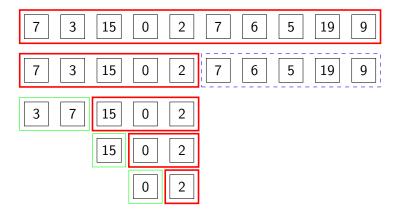


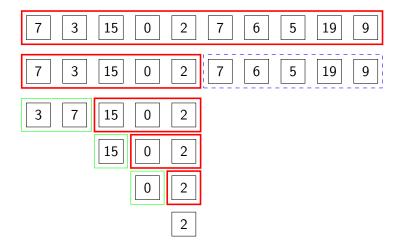


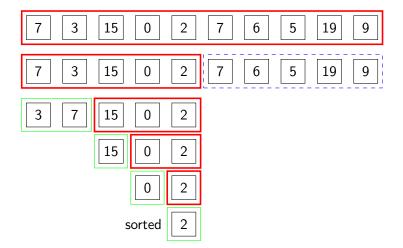


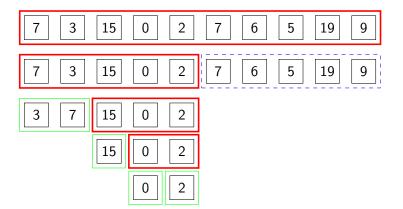


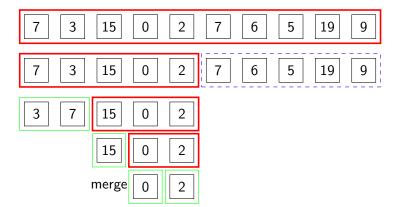


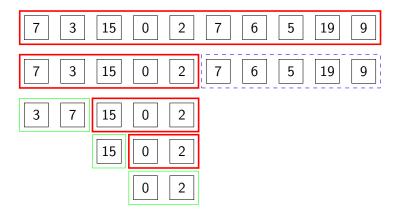


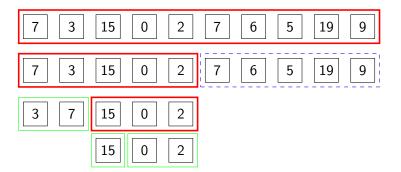


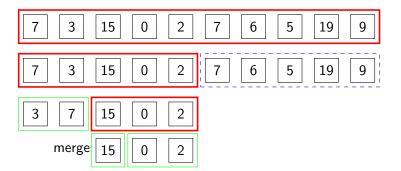


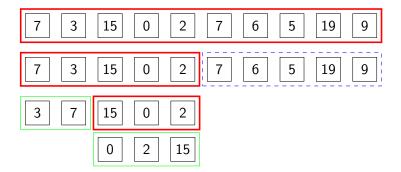


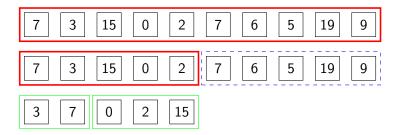


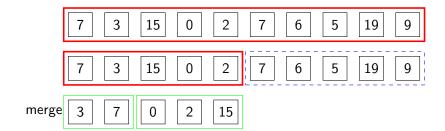


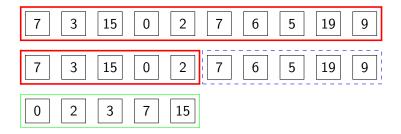


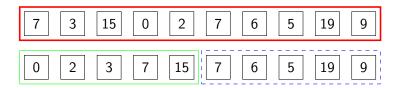


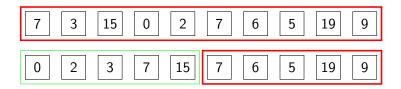


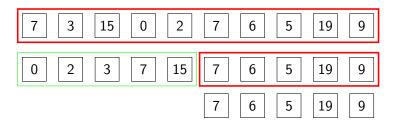


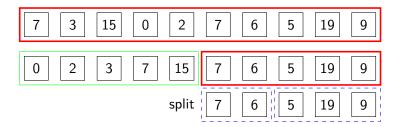


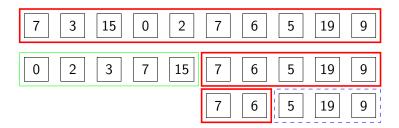


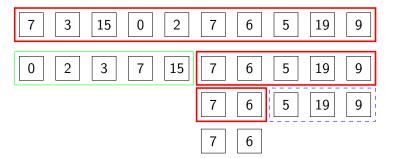


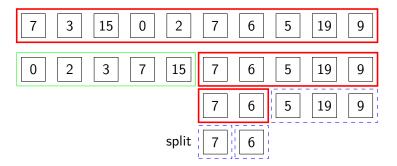


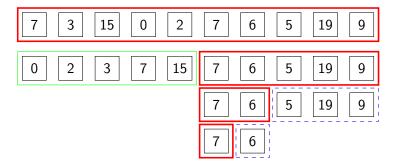


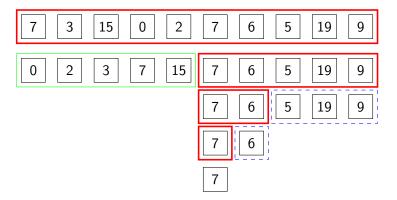


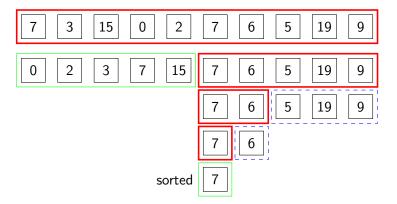


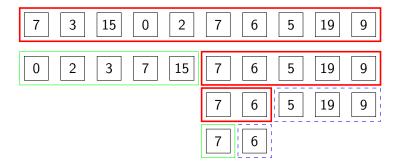


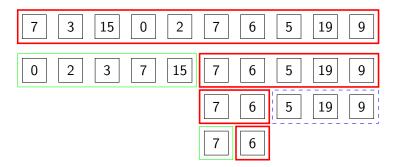


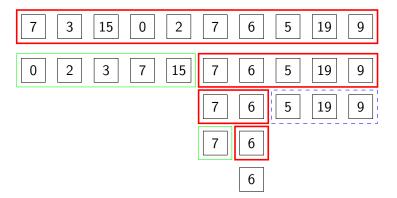


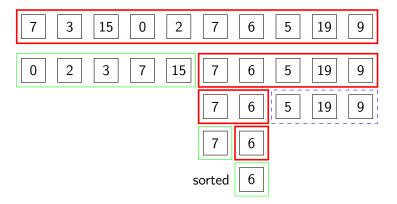


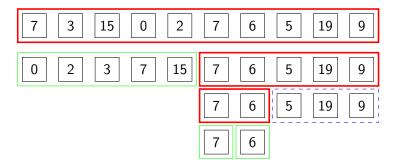


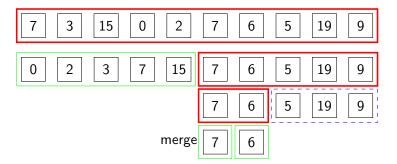


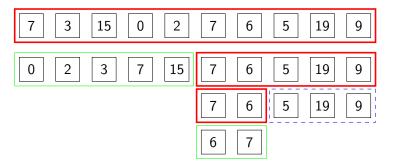


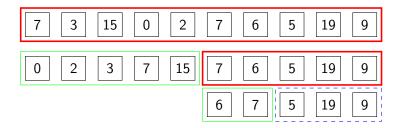


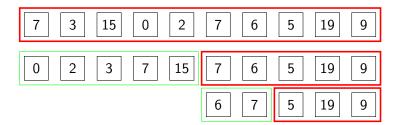


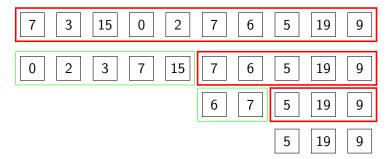


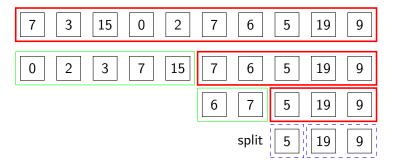


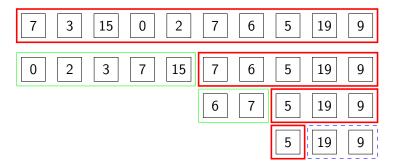


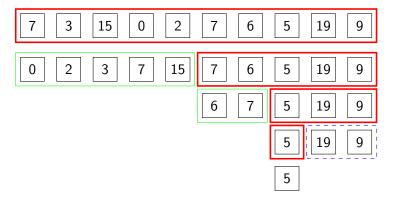


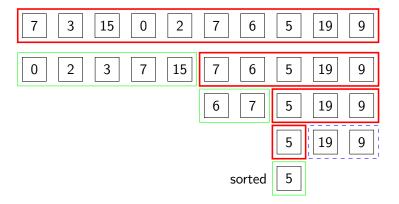


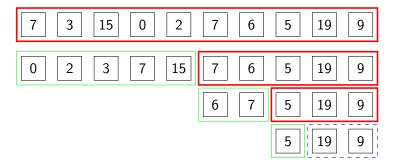


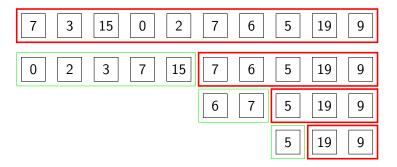


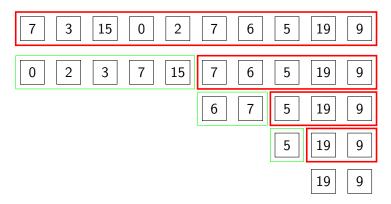


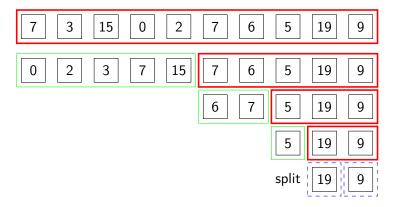


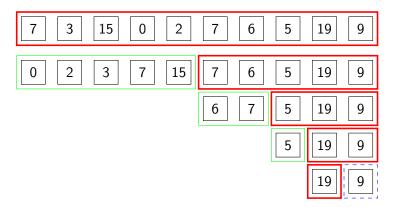


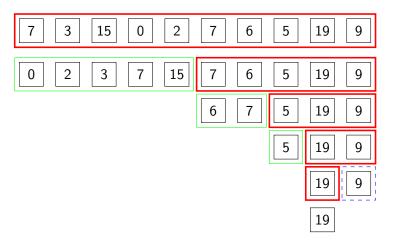


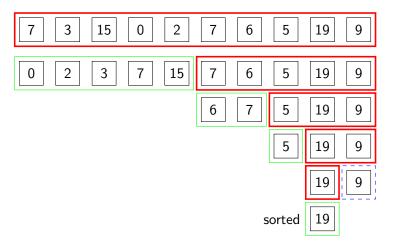


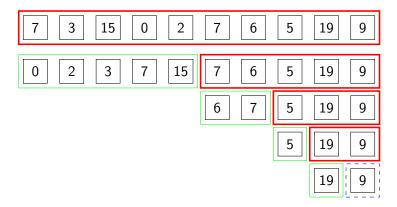


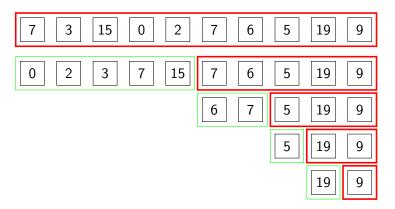


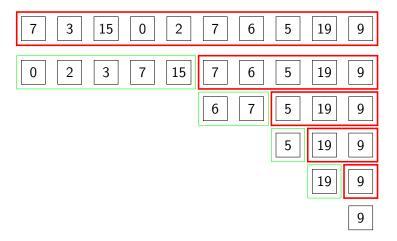


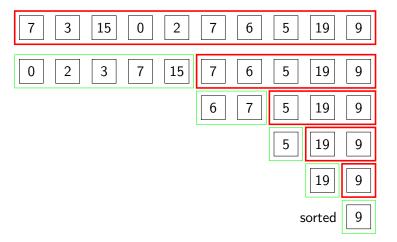


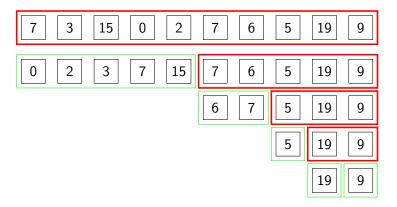


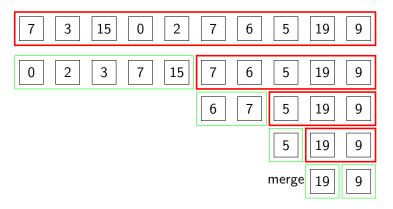


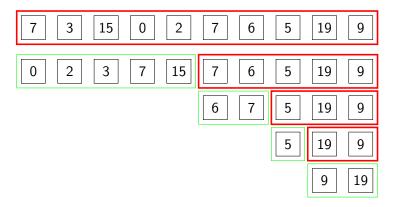


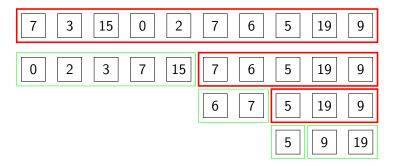


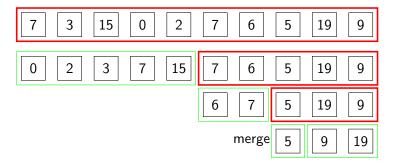


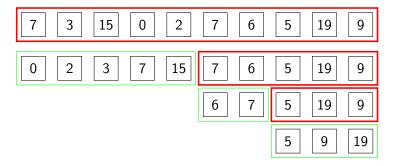


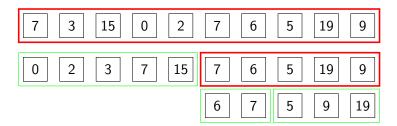


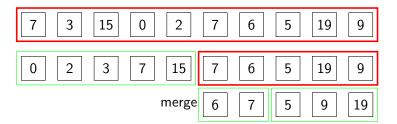


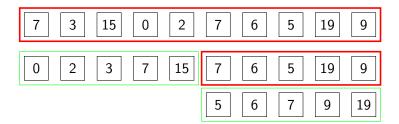


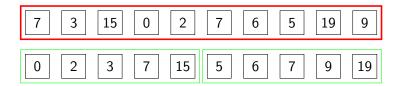


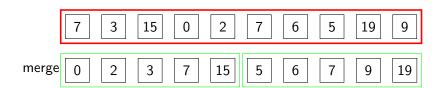














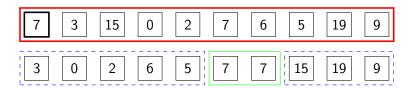
0 2 3 5 6 7 7 9 15 19

Quicksort

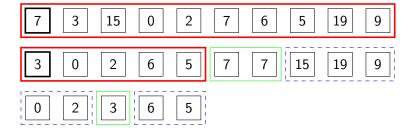
- ► Algoritm:
 - Välj ut ett pivåelement.
 - ▶ Dela upp listan i tre delar: Less, Equal, Greater.
 - Sortera Less och Greater rekursivt.
 - ► Slå ihop Less+Equal+Greater.

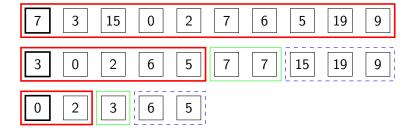
7 3 15 0 2 7 6 5 19 9

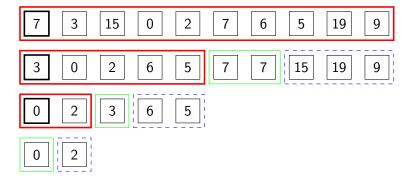
7 3 15 0 2 7 6 5 19 9

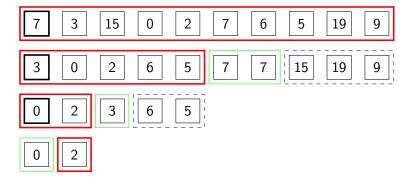


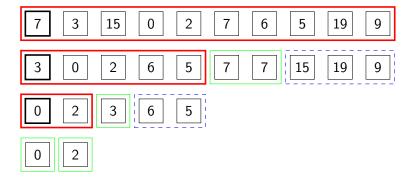


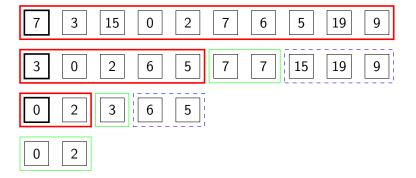


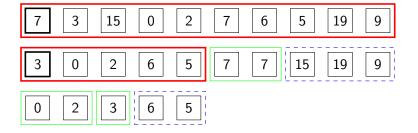


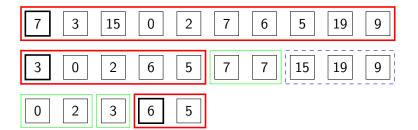


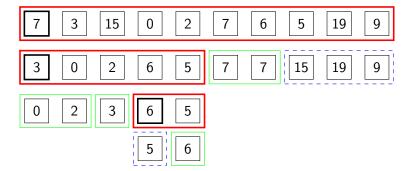


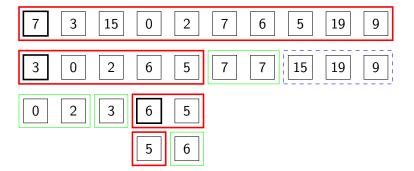


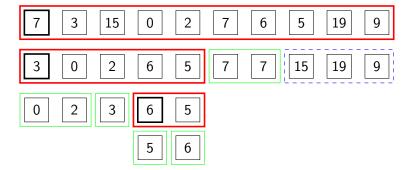


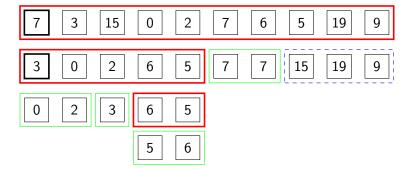


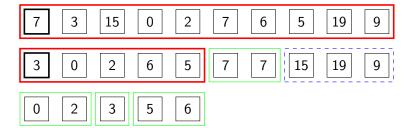


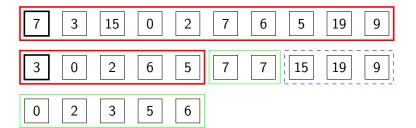


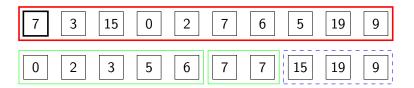


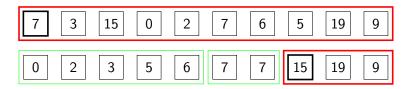


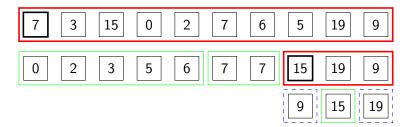


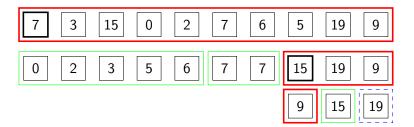


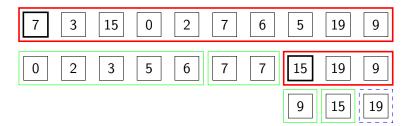


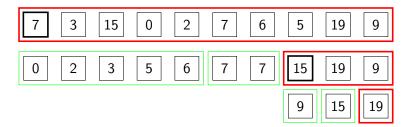


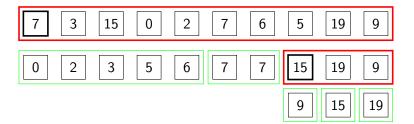


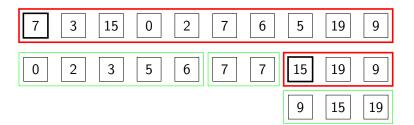


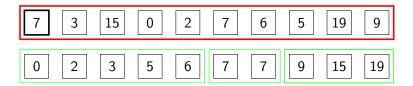


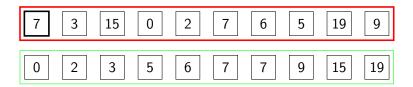












0 2 3 5 6 7 7 9 15 19

Quicksort — val av pivåelement

- \triangleright $O(n \log n)$ i bästa fallet.
- ► Valet av pivåelement kritiskt:
 - ▶ Vill ha ett pivåelement som har ett mitten-värde.
 - Vid sned fördelning får man insticks-/urvalssortering med $O(n^2)$.
- ► Alternativ (eftersträvar en enkel tilldelning):
 - ► Välj första/sista, slumpmässigt.
 - Medel/median mellan några stycken.
 - Största av de två första som skiljer sig åt.

Några algoritmer sammanfattade

Algoritm	Tidskomplexitet	Stabil?	Minnesbehov
Insertion Sort	$O(n^2)$	Ja	O(1)
Selection Sort	$O(n^2)$	Nej^1	O(1)
Bubble Sort	$O(n^2)$	Ja	O(1)
Merge Sort	$O(n \log n)$	Ja	O(n)
Quicksort	$O(n \log n) - O(n^2)$	Nej	$O(\log n) - O(n)$
Heapsort	$O(n \log n)$	Nej	O(1)
(Radix Exchange Sort)	O(wn)	Nej	O(w)

¹Ja om O(n) extra minne.