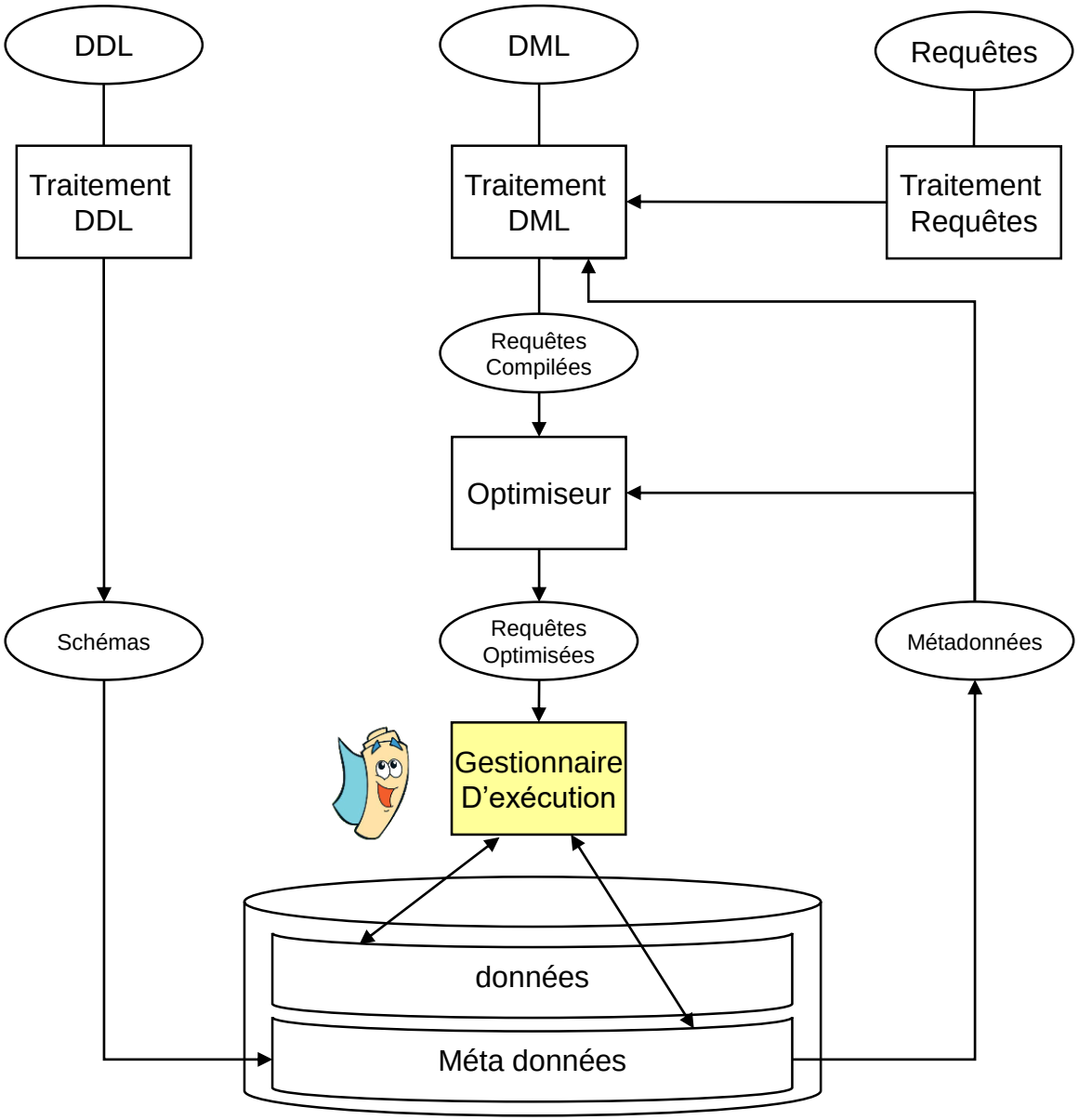
Architectures des Systèmes de Bases de Données

Merge Join without duplicates

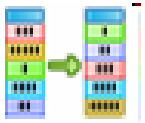


D'après C.J DATE

DDL : langage de définition des données; DML : langage de manipulation des données

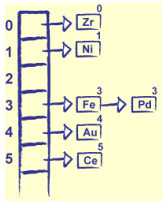
Memory join algorithms

- Nested loop



- Merge join

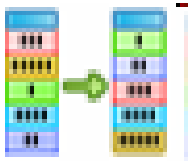
- Hash join

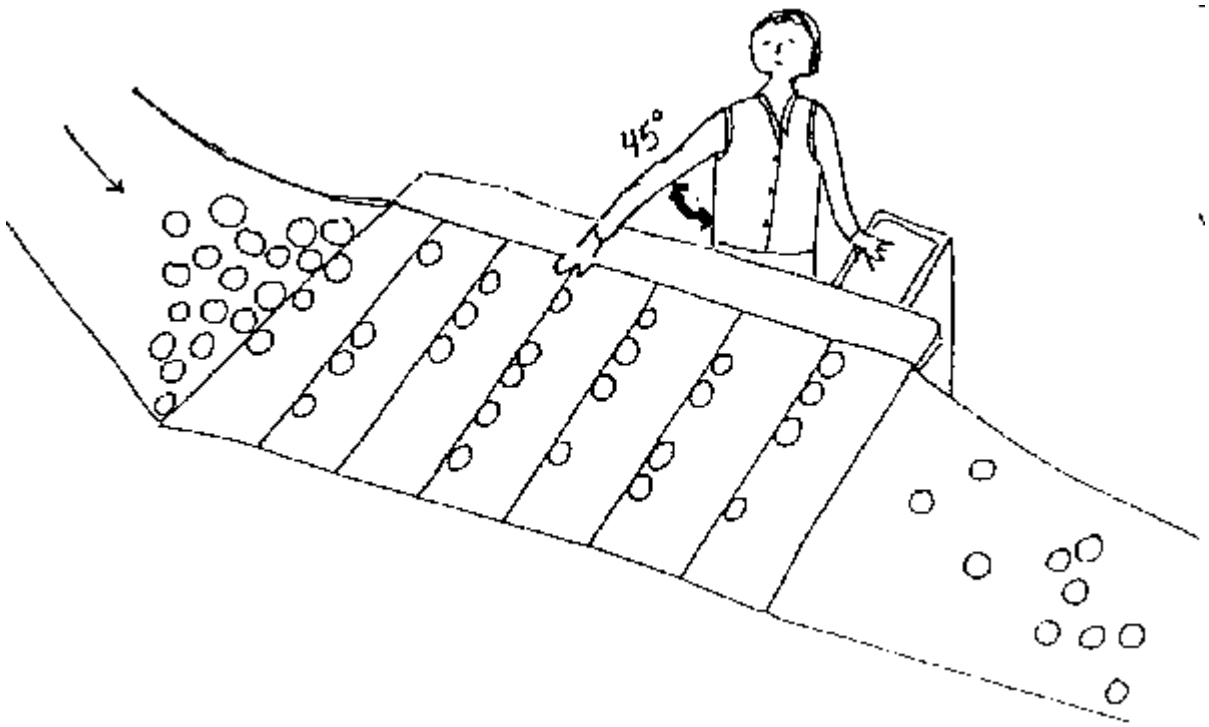


Où sont les problèmes de performances ?

- Dans les opérations de comparaisons (Tests).
- NestedLoop (boucles imbriquées)
  - Complexité O(n2)
- Optimisation :
  - Comment réduire le nombre de tests ?

Il faut trier !!!!





Merge Join

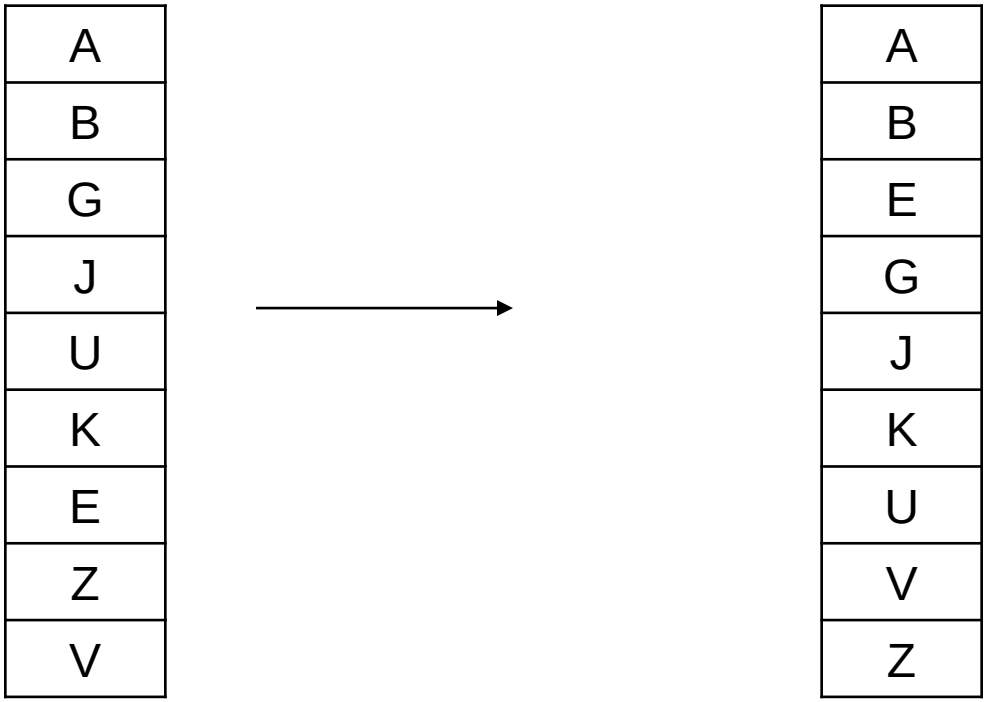
- Quand les tableaux sont triés l'opération de jointure s'appelle:

Merge Join

Principe

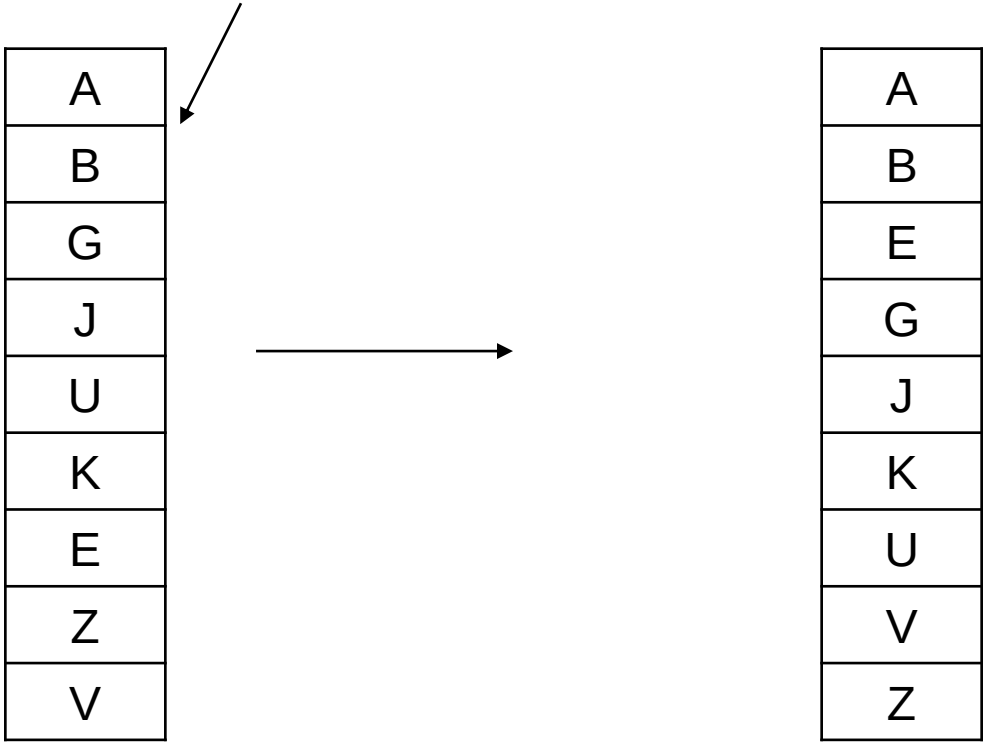
- 1) trier les tableaux
- 2) fusionner les tableaux

R Triée



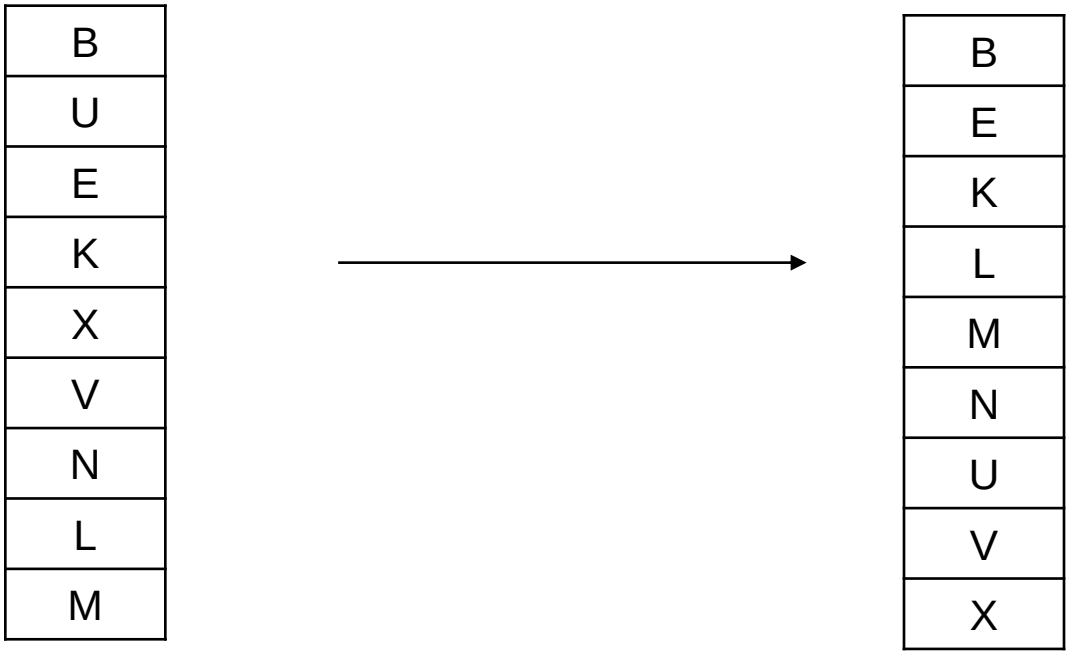
R Triée

no duplicates



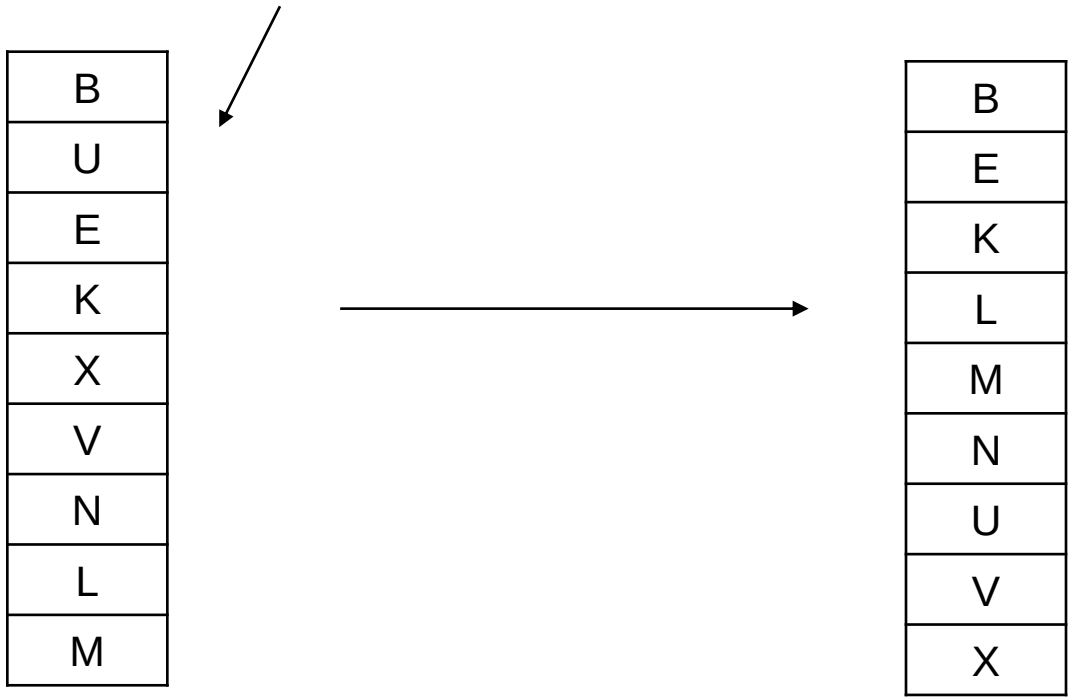
without duplicates

S triée



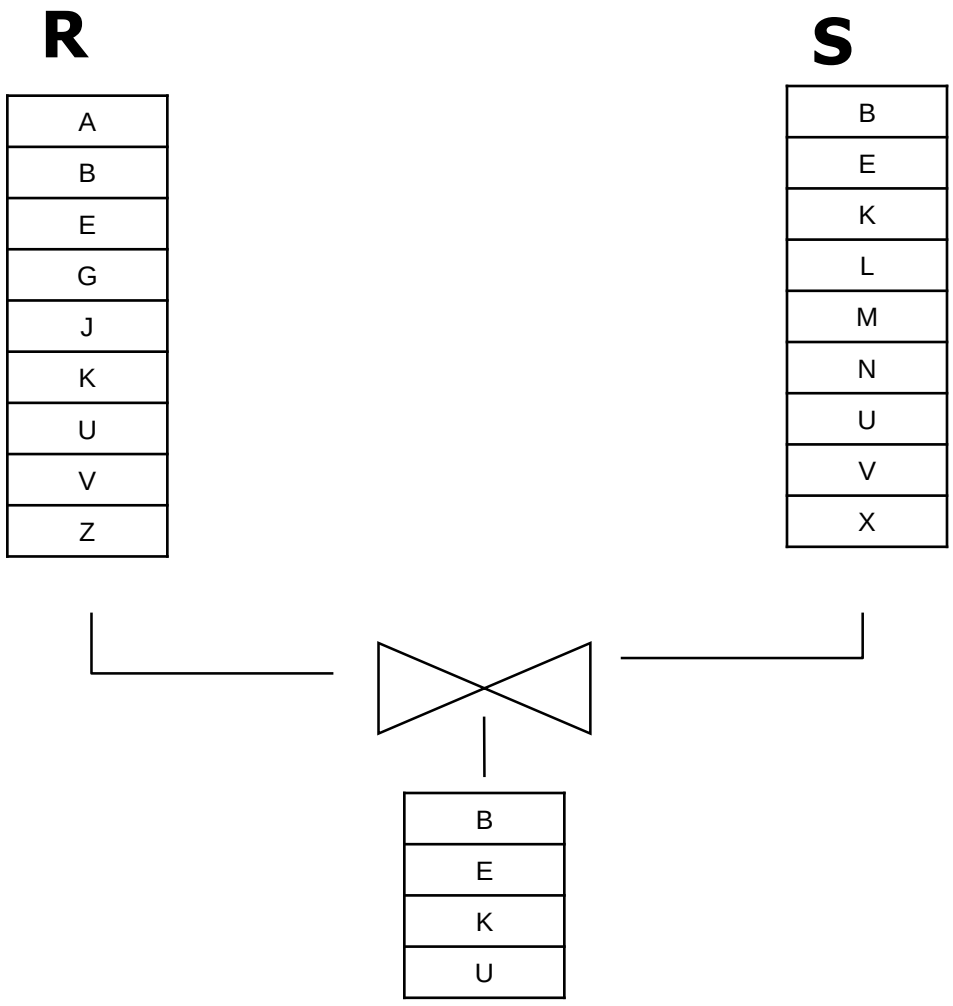
S triée

no duplicates

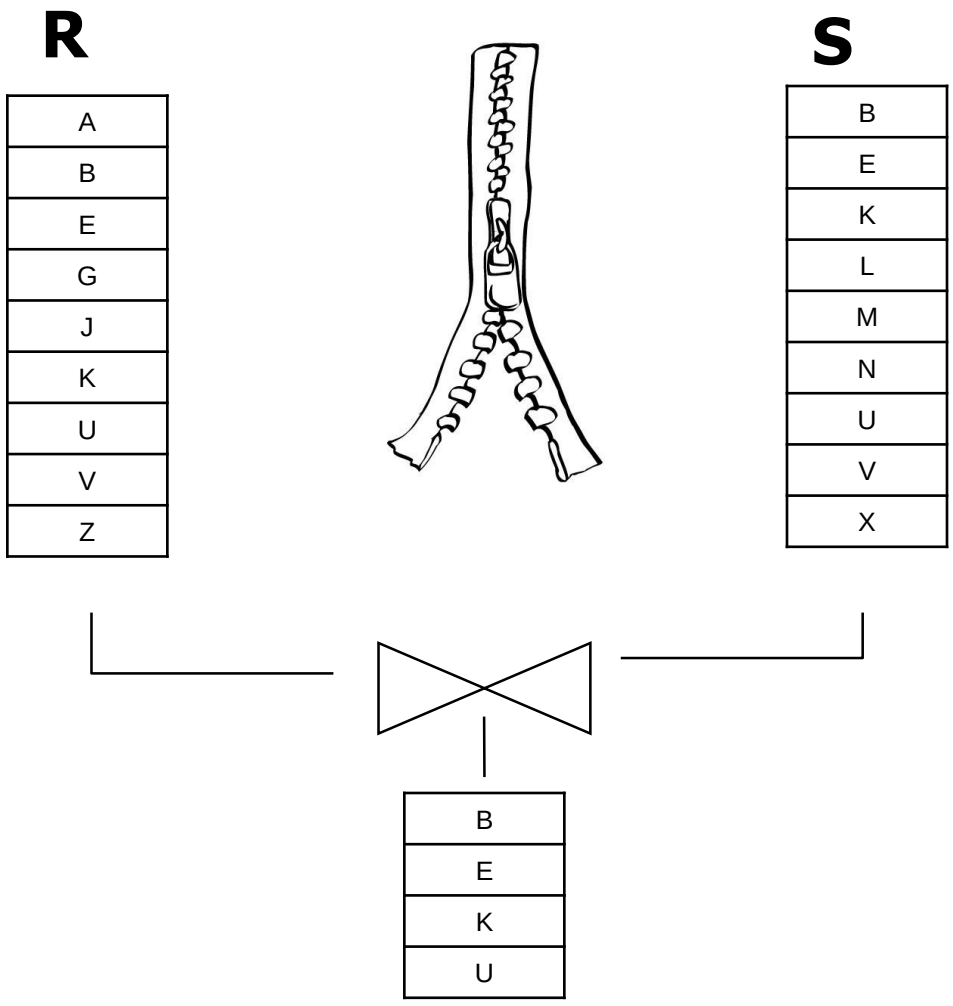


without duplicates

Résultat de la jointure



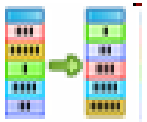
Résultat de la jointure



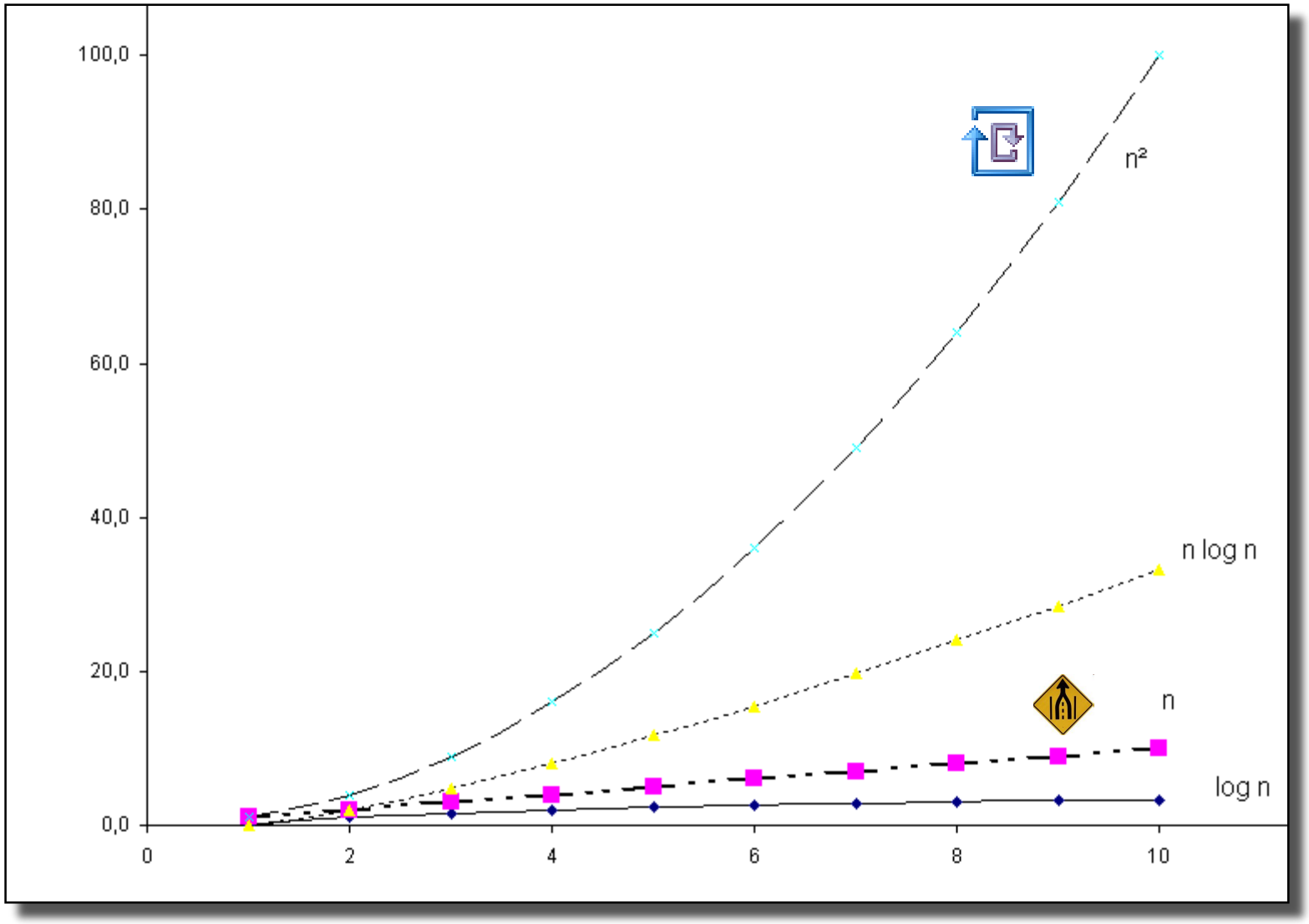
Complexity

- NestedLoop = R*S

- MergeJoin = R+S
  - Plus sort complexity.

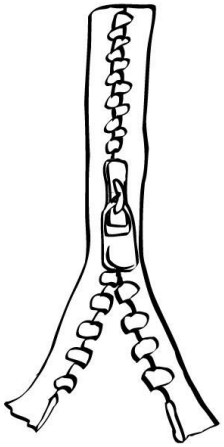


Complexité



A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X



B
E
K
U
V

without duplicates

Merge Join

Algorithme de Jointure sans doublon

Algo2

```
i=j=1

while  i < (#tuples in R) and j < (#tuples in S)
|
|       if R(i) = S(j) then
|       |               output(R(i), S(j))
|       |               i=i+1
|       endif
|
|       else if R(i) > S(j) then j=j+1
|
|       else R(i) < S(j) then i=i+1
|
end while
```

Merge Join

R Triée

A
B
G
J
U
K
E
Z
V

A
B
E
G
J
K
U
V
Z

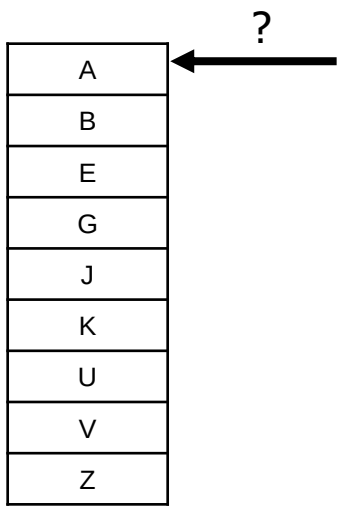
without duplicates

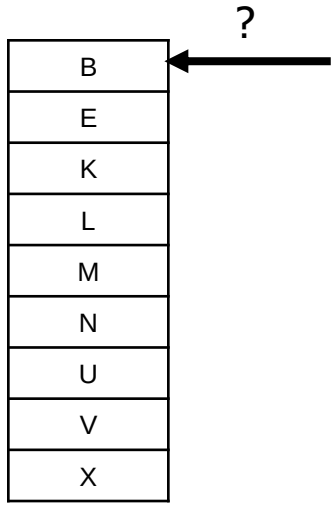
S triée

B
U
E
K
X
V
N
L
M

B
E
K
L
M
N
U
V
X

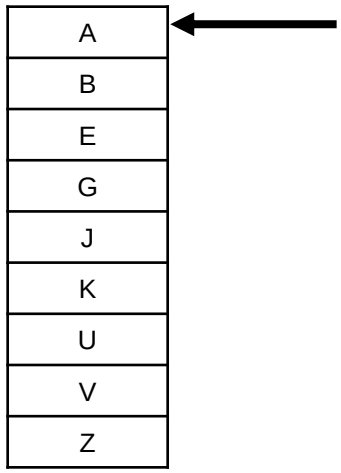
without duplicates

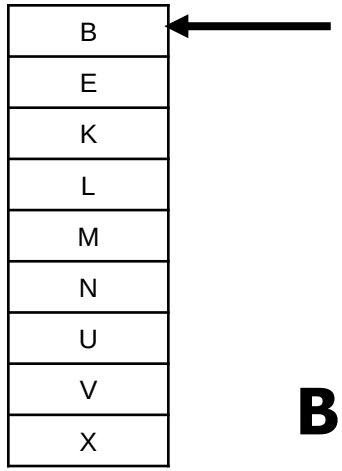




without duplicates

Merge Join

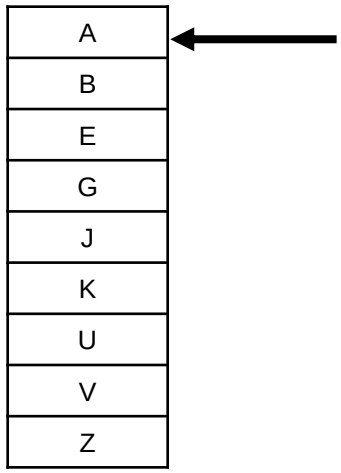


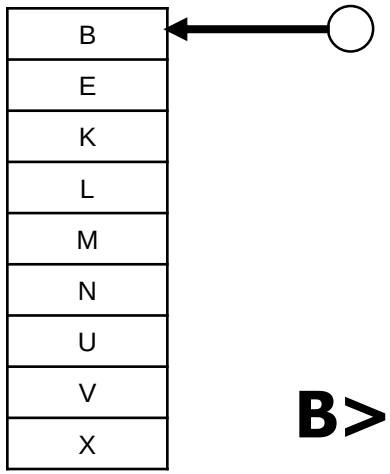


B>A

without duplicates

Merge Join

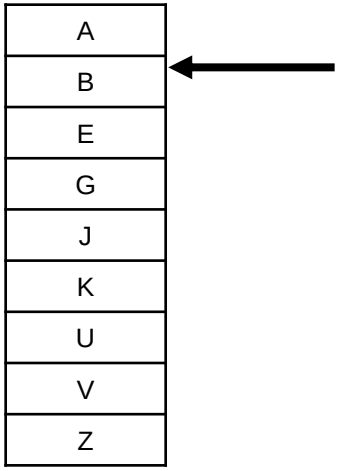


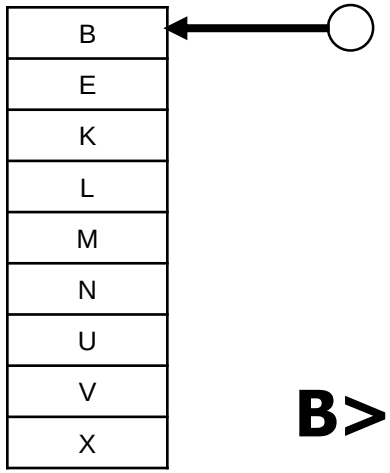


B>A

without duplicates

Merge Join

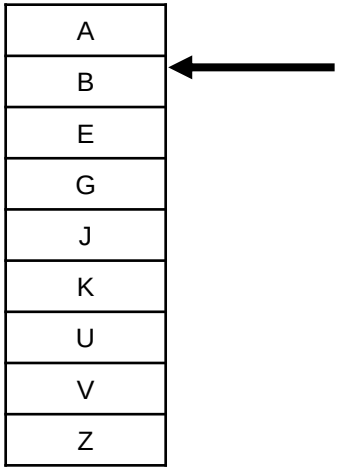


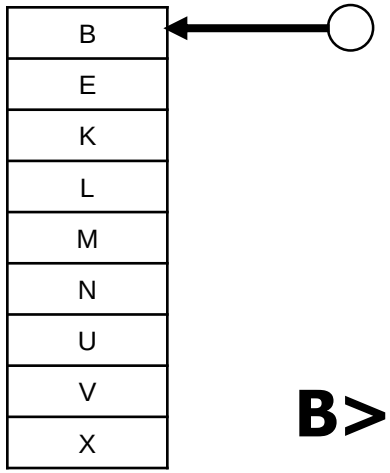


B>A

without duplicates

Merge Join

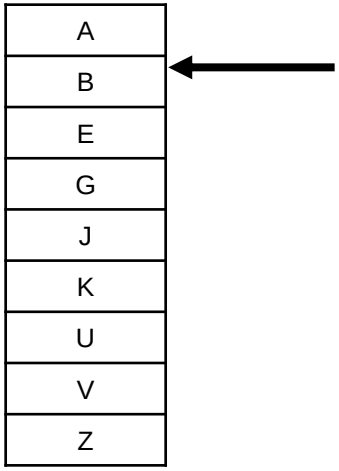


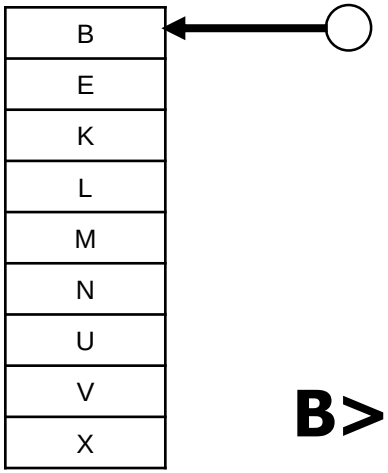


B>A

without duplicates

Merge Join



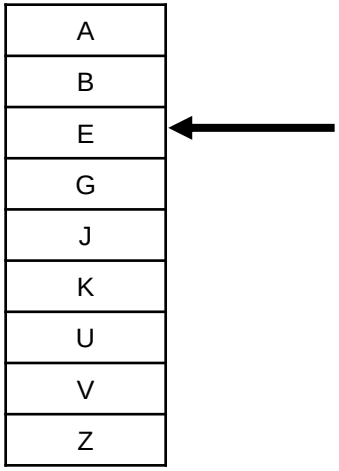


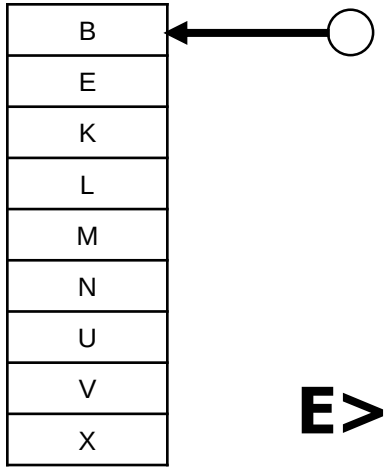
B>A

Match

without duplicates

Merge Join

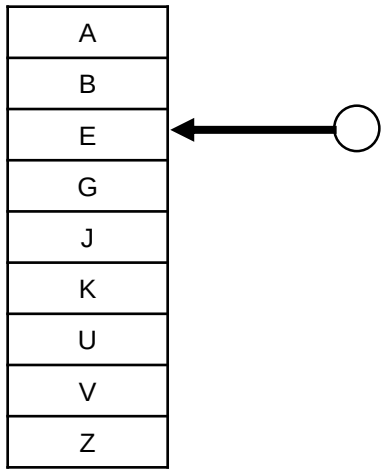


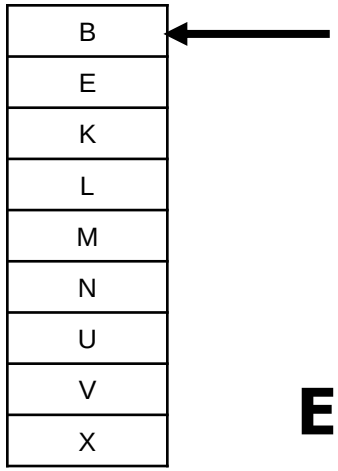


E>B

without duplicates

Merge Join

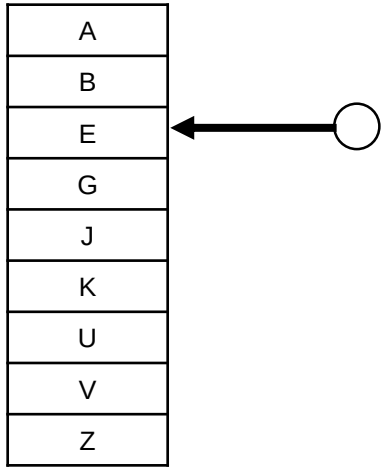


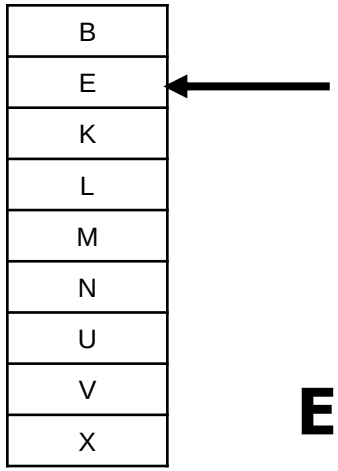


E>B

without duplicates

Merge Join





E>B

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

Match

B
E

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

G>E

B
E

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

K>G

B
E

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

K>G

B
E

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

G<K

B
E

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

J<K

B
E

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

Match

B
E
K

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

U>K

B
E
K

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

U>K

B
E
K

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

U>L

B
E
K

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

U>M

B
E
K

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

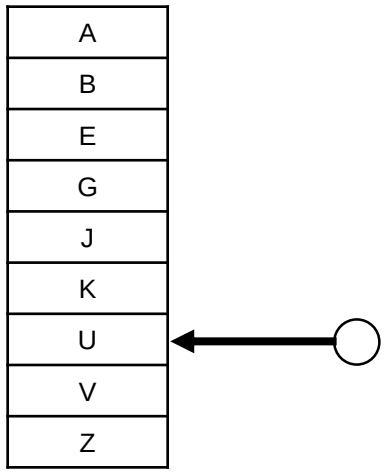
B
E
K
L
M
N
U
V
X

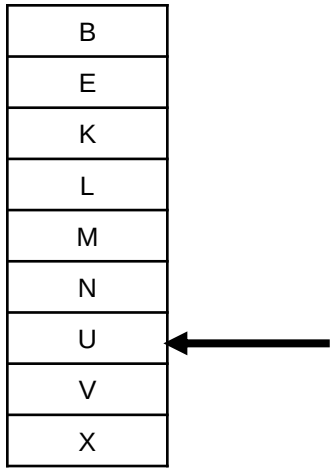
U>N

B
E
K

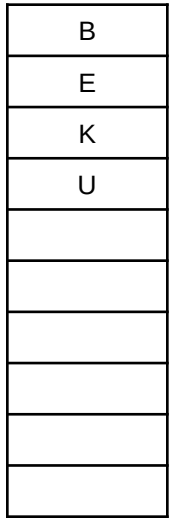
without duplicates

Merge Join





Match



without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

B
E
K
L
M
N
U
V
X

V>U

B
E
K
U

without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

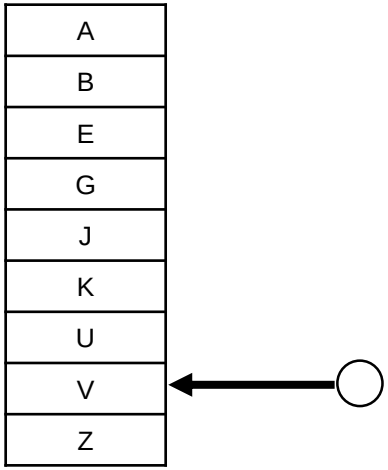
B
E
K
L
M
N
U
V
X

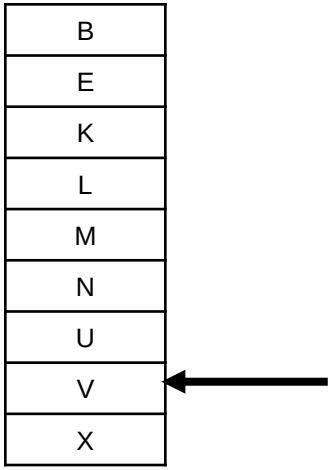
V>U

B
E
K
U

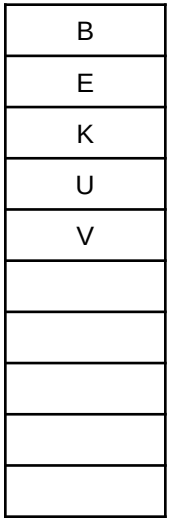
without duplicates

Merge Join



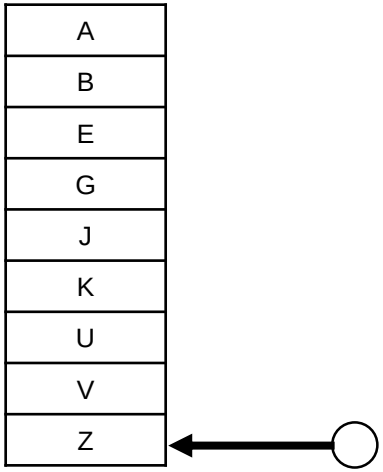


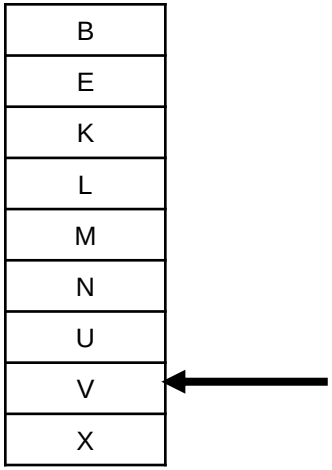
Match



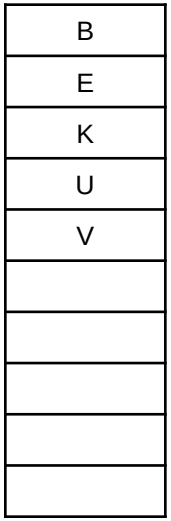
without duplicates

Merge Join





Z>V



without duplicates

Merge Join

A
B
E
G
J
K
U
V
Z

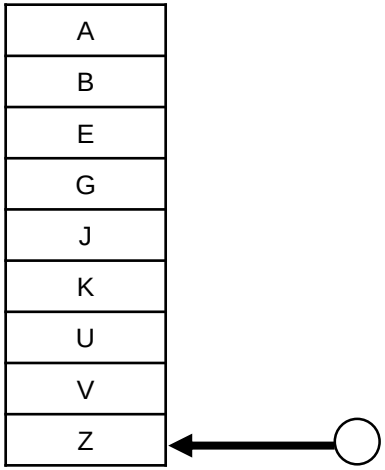
B
E
K
L
M
N
U
V
X

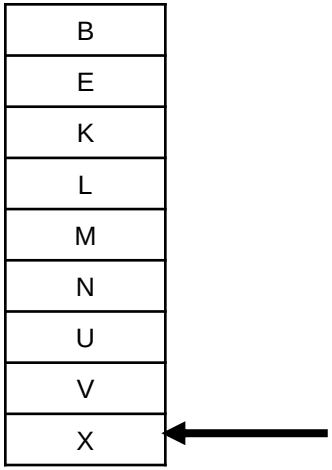
Z>X

B
E
K
U
V

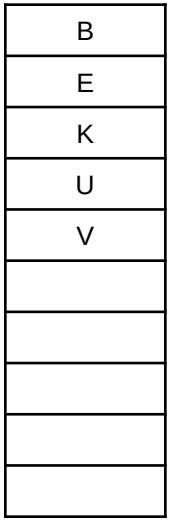
without duplicates

Merge Join





Fin



without duplicates

Merge Join