



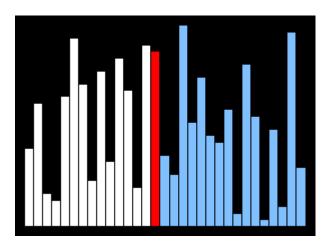
Algoritmos y Estructuras de Datos I

Facultad de Ciencias Exactas y Tecnología
Universidad Nacional de Tucumán
2023

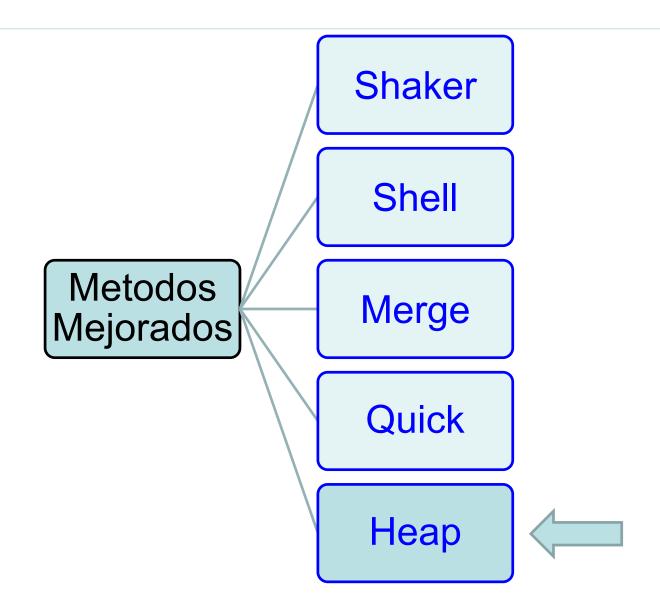




Ordenación(3)



Métodos de Ordenación Interna



MÉTODO del montículo o HEAPSORT

En 1964 J.W.J. Williams presenta este método cuya filosofía es la del método de selección directa, en el sentido de que se encuentra el mayor de una colección de n claves, luego la siguiente mayor y así.

En base a las operaciones de la estructura de datos Cola de Prioridad (PQ) se puede definir este método de manera elegante y eficiente.

La idea es simple, armar una PQ con los elementos a ordenar, luego sacarlos de PQ y reubicarlos en el mismo arreglo.

El método se puede implementar de manera que no use memoria extra, mejora propuesta por R. W. Floyd en ese mismo año.

Cola de prioridad (priority queue) PQ(ITEM)

Una *Cola de Prioridad* es una fila en la que cada ítem tiene asociada una prioridad con una relación de orden >.

La operación de selección siempre elige el elemento de mayor prioridad.

Especificación algebraica - OPERACIONES:

A) Sintaxis:

PQVACIA : → PQ

ENPQ: PQ X ITEM → PQ

ESPQVACIA: PQ → BOOLEAN

MAYOR : PQ → ITEM U {indefinido}

DEPQ: PQ → PQ

Cola de prioridad PQ(ITEM)

Especificación algebraica - OPERACIONES:

B) Semántica: $\forall q \in PQ$, $\forall i \in ITEM$

```
ESPQVACIA(PQVACIA) \equiv TRUE

ESPQVACIA(ENPQ(q,i)) \equiv FALSE
```

```
MAYOR(PQVACIA) ≡ indefinido

MAYOR(ENPQ(q,i)) ≡ si ESPQVACIA(q) entonces i

sino si i>MAYOR(q) entonces i

sino MAYOR(q)
```

DEPQ(PQVACIA) ≡ PQVACIA

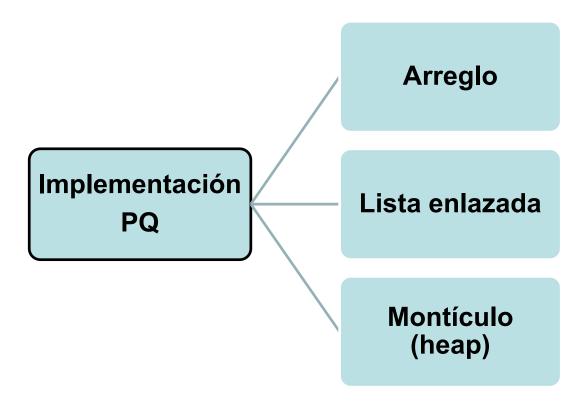
DEPQ(ENPQ(q,i)) ≡ si ESPQVACIA(q) entonces PQVACIA

sino si i>MAYOR(q) entonces q

sino ENPQ(DEPQ(q),i)

PQ(ITEM) Implementación

Se puede realizar la Implementación de PQ con:



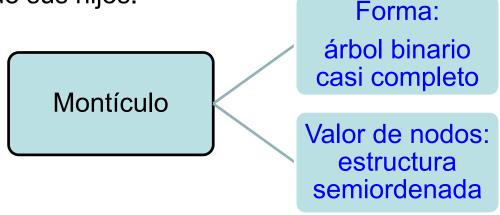
Definición de Montículo

Un *Montículo* es un árbol que cumple:

• Es un árbol binario casi completo, en el penúltimo nivel tiene todos los nodos, en el ultimo nivel las hojas se ubican de izquierda a derecha.

• Es una estructura *semiordenada*, el ítem almacenado en cada

nodo es mayor que sus hijos.

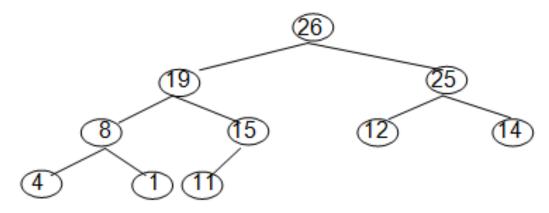


Un montículo se implementa con un arreglo que es la estructura ideal para albergar los ab casi completos.

Implementación de un Montículo

Los arboles binarios casi completos se pueden representar secuencialmente en un arreglo.

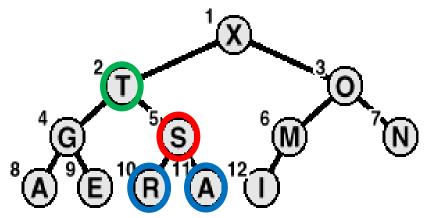
- la raíz se ubica en el índice 1
- sus hijos en las posiciones 2 y 3
- el nivel siguientes en las posiciones 4, 5, 6 y 7 Por Ejemplo:



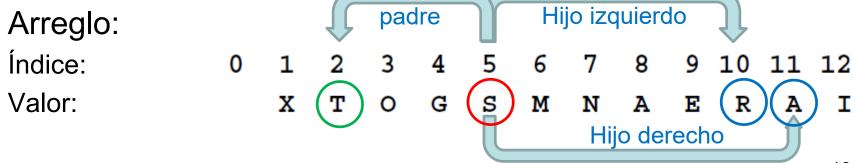
A = [26, 19, 25, 8, 15, 12, 14, 4, 1, 11]

Implementación de un Montículo

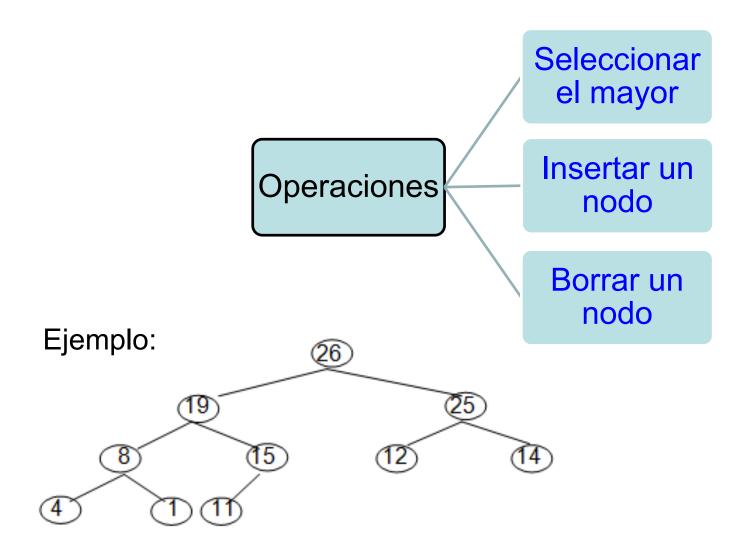
Ejemplo:



	Índice del arreglo
Raíz(árbol)	1
Padre(nodo(i))	i/2
Hijolzq(nodo(i))	2*i
HijoDer(nodo(i))	2*i+1



Operaciones para Montículo



Operaciones para Montículo

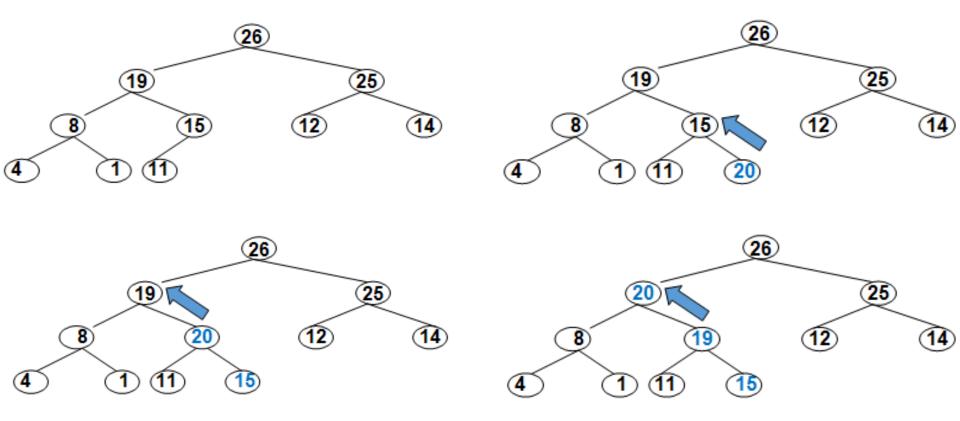
INSERTAR un item: la inserción de un nuevo nodo debe mantener la condición de heap, debe seguir siendo un ab casi completo y semiordenado.

El proceso es el siguiente:

- Si el heap tiene n nodos, se inserta en la primera componente libre del arreglo, la n+1-ésima.
- Si viola con su valor la condición de heap, se lo intercambia con su padre. Y así se sigue con los intercambios hasta que ocupe su lugar sin violar la condición de orden o hasta llegar a la raíz del árbol.

El algoritmo para que el ítem suba a su posición final es Arriba:

Ejemplo Algoritmo Arriba: insertar 20



Algoritmo Arriba (A, k)

ENTRADA:

A ∈ ARREGLO(item), de tamaño n

k: índice de A donde está ubicado el item.

SALIDA:

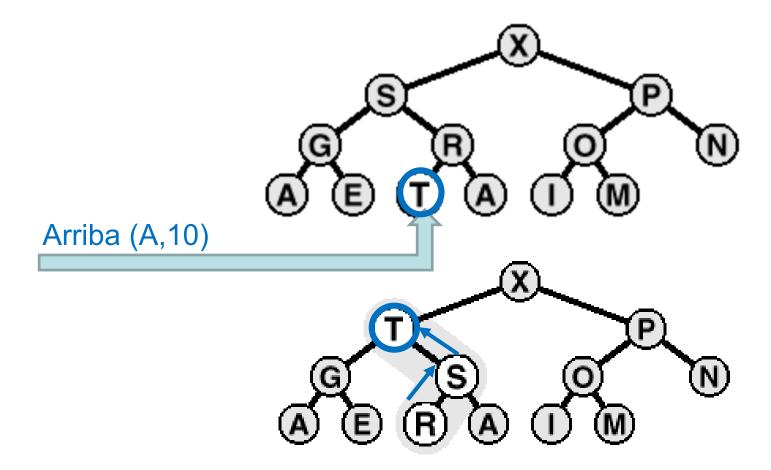
A ∈ ARREGLO(item), con todos sus elementos en condición de heap Auxiliares: aux ∈ item

```
\begin{aligned} \text{aux} &\leftarrow \mathsf{A}(\mathsf{k}); \qquad \text{$''$ se guarda el valor del nodo que tiene que subir en el árbol} \\ \mathsf{MIENTRAS} \; (\mathsf{k} > 1) \; \mathsf{AND} \; (\mathsf{A}(\mathsf{k} \, / \, 2) \leq \mathsf{aux} \;) \; \mathsf{HACER} \; \text{$''$ tiene padre y es menor que el hijo} \\ \mathsf{A}(\mathsf{k}) &\leftarrow \mathsf{A}(\mathsf{k} \, / \, 2) \; \text{$''$ se reemplaza al hijo por su padre} \\ \mathsf{k} &\leftarrow \mathsf{k} \, / \, 2 \; \text{$''$ se sube un nivel en el arbol} \\ \mathsf{A}(\mathsf{k}) &\leftarrow \mathsf{aux}; \; \text{$''$ se coloca al nodo en su posición final} \end{aligned}
```

FIN

Peor caso para este algoritmo es O(log₂n), donde n es el nro. de ítems del arreglo.

Ejemplo: aplicación de Arriba:



Operaciones para Montículo

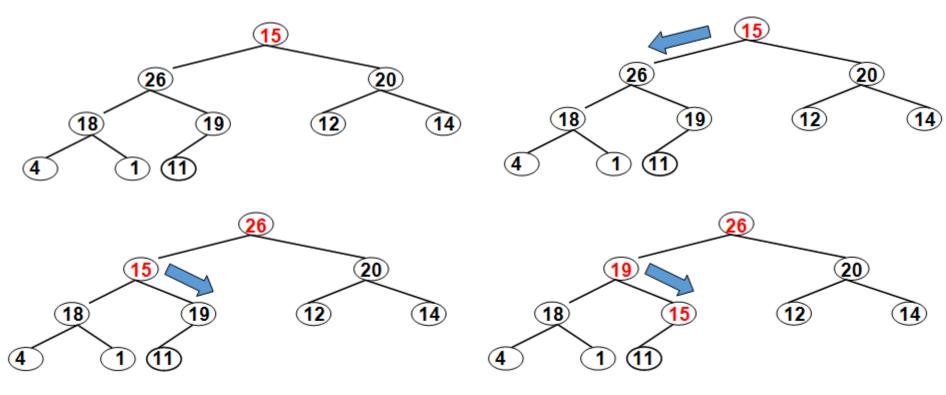
BORRAR un item: la operación más frecuente es borrar el item de mayor valor del heap, esto es el que está en la raíz del árbol.

El procedimiento es el siguiente:

- Para borrar el mayor elemento, primero se lo sustituye por el nodo ce la última hoja que pasa a ser la raíz.
- Para restablecer la condición de heap si es que no se cumple que sea mayor que sus hijos, se lo intercambia con el mayor de sus hijos, y así se sigue con los intercambios mientras corresponda o hasta que se llegue a una hoja.

El algoritmo para que el ítem en la raíz baje a su posición definitiva es Abajo:

Ejemplo Algoritmo Abajo



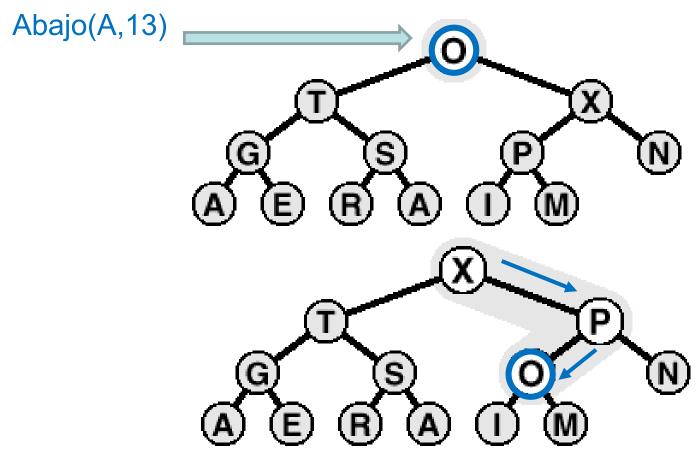
```
Algoritmo Abajo (A,tamanio)
ENTRADA: A ∈ ARREGLO(item), tamanio: tamaño real de A
SALIDA: A ∈ ARREGLO(item), con todos sus elementos en condición de heap
Auxiliares: aux \in item, sigue \in bool
k \leftarrow 1; aux \leftarrow A(k); sigue \leftarrow true
MIENTRAS ( k \le tamanio/2 AND sigue ) HACER // tiene al menos un hijo
    j \leftarrow 2^*k; // hijo izquierdo
     SI j<tamanio AND A(j) \leq A(j+1) ENTONCES // \exists el hijo derecho y es mayor que el izq.
          j ← j+1
                                          // en índice j esta el mayor de sus hijos
     SI aux \geq A(j) ENTONCES
                                          // aux ya es mayor que sus hijos
          sigue ← false
                                          // no itera mas
     SINO
          A(k) \leftarrow A(j)
                                          // reemplaza al padre por el mayor de sus hijos
          k ← j
                                          // se baja un nivel en el árbol
     FIN SI
                                               Peor caso para este algoritmo es
FIN MIENTRAS
                                               O(log_2n), donde n es el nro. de ítems.
```

18

 $A(k) \leftarrow aux // se coloca al nodo en su posición final$

FIN

Ejemplo: aplicación de Abajo:



Operaciones para PQ

Costo de las operaciones de PQ implementada con Monticulo:

Si la PQ tiene n items:

- ENPQ $\epsilon O(\log_2 n)$
- ESPQVACIA € O(1)
- MAYOR ϵ O(1)
- DEPQ $\epsilon O(\log_2 n)$

MÉTODO HEAPSORT

Algoritmo Heap (A,n)

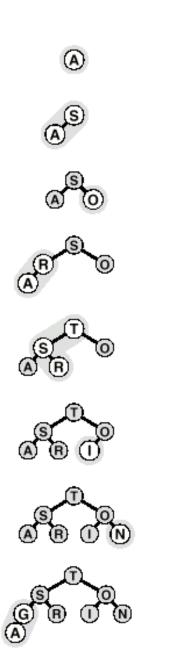
```
Entrada: A: arreglo de (1.. MAX) elementos de tipoitem
             n: nro. entero de datos a ordenar
Salida: A arreglo ordenado
Auxiliares: pq ∈ tipopq; x∈tipoitem; i ∈ entero
pq \leftarrow pqvacia; \longrightarrow \epsilon O(1)
PARA i=1,n HACER \longrightarrow nit=n
enpq(pq,A(i)) \longrightarrow \in O(log<sub>2</sub>n)
                                                              € O(nlog<sub>2</sub>n)
PARA i=n,1 paso (-1) HACER \longrightarrow nit=n
                                                                             T(n) ε O(nlog₂n)
          x \leftarrow mayor(pq) \longrightarrow \epsilon O(1)
                                                              _€ O(nlog<sub>2</sub>n)
          depq(pq) \longrightarrow \epsilon O(log_2 n)
          A(i) \leftarrow x \longrightarrow \epsilon O(1)
FIN
```

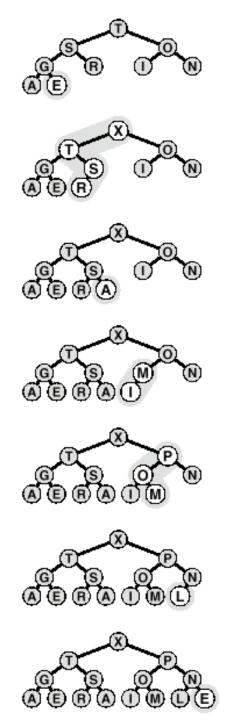
HEAPSORT

Ejemplo:. Ordenar con HeapSort la cadena: **ASORTINGEXAMPLE**

PASO 1)

Construcción de un heap (enpq) con los datos a ordenar:





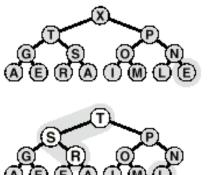


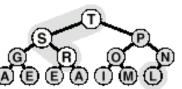
HEAPSORT

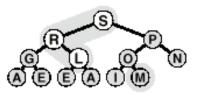
Ejemplo:

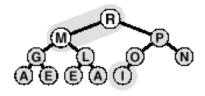
PASO 2)

Desarmado (depq) para ordenar según el heap: **XTSRPONMLIGEEAA**

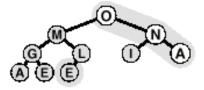


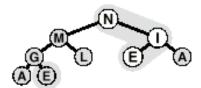


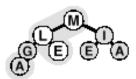


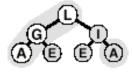


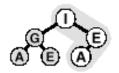


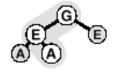


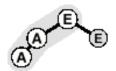




















Propiedades Algoritmo HEAP

Tiempo de Ejecución: Por la forma de implementación sobre una cola de prioridad Heapsort requiere tiempo proporcional a O(n.log₂n) aun en el caso más desfavorable.

$$T(n) \in O(n \cdot \log_2 n)$$
.

No es claro en que casos puede esperarse el mejor o peor rendimiento. Pero generalmente el método parece funcionar un poco mejor con secuencias iniciales en orden inverso.

Memoria: si se programa para que el arreglo y el montículo compartan el almacenamiento no necesita espacio extra.

Sensibilidad: nada sensible.

Estabilidad: NO es estable

TIEMPO DE EJECUCION DE DISTINTOS ALGORITMOS DE ORDENACIÓN

La tabla proporcionada por Wirth muestra el tiempo de ejecución en milisegundos consumido por distintos métodos. Las tres columnas corresponden a un arreglo ordenado, una permutación aleatoria y el arreglo ordenado en sentido inverso, todos del mismo tamaño. Los datos que se utilizan están formados sólo por una clave.

METODO	Ordenado	Aleatorio	Inverso
SELECCION DIRECTA	1907	1956	2675
INSERCION DIRECTA	23	1444	2836
INTERCAMBIO DIRECTO	2165	4054	5931
SHAKER-SORT	9	3642	6520
SHELL-SORT	116	349	492
QUICK-SORT	69	146	79
MERGE-SORT	234	242	232
HEAP-SORT	253	241	226

ALGORITMOS DE ORDENACIÓN Demos

Muy bueno:

http://www.sorting-algorithms.com/

50+ Sorts, Visualized - Scatter Plot:

https://www.youtube.com/watch?v=DSMCZZGbZo4

50+ Sorts, Visualized - Reversed Inputs:

https://www.youtube.com/watch?v=qtRU2Xn76Bc

50+ Sorts, Visualized - Almost Sorted Inputs:

https://www.youtube.com/watch?v=UdTyfJ4zJDA

METODOS LINEALES

Todos los métodos de ordenación presentados sirven para clasificar ítems con cualquier tipo de campo clave, siempre que en el mismo esté definido una relación de orden. Los algoritmos se definen en términos de operaciones básicas de comparaciones y movimientos.

Se ha demostrado que son necesarios O(nlog₂n) pasos para ordenar n elementos siempre que sea un método de propósito general para cualquier tipo de claves.

Los métodos con costo lineal no son aplicables a cualquier caso sino a un rango de problemas bien definidos. Sirven para un tipo particular de clave tal que tengan condiciones especiales como: ser números enteros y pertenecer a un cierto rango.

METODOS LINEALES

El método Radix tiene su origen en 1887 cuando se comenzó a usar por Herman Hollerith en las maquinas tabuladoras. En 1923 los algoritmos Radix se popularizaron con las tarjetas perforadas. Recién se programó para una computadora en 1954 (Harold Seward, MIT).

Actualmente Radix sort se usa frecuentemente para ordenar enteros y cadenas de bits. Se ha probado que es mucho mas rápido que cualquier otro metodo de propósito general.

Se puede implementar en sus 2 versiones:

- LSD (least significant digit): comienza ordenando las unidades.
- MSD (most significant digit):comienza con el dígito mas significativo.

Método Radix : permite clasificar n claves numéricas de M dígitos en cualquier base B.

EJEMPLO: Arreglo a ordenar:

A=[036 909 100 025 101 049 064 471 381 004 555 186 100 038 671 001 010 999 899 002]

Son n=20 claves numéricas de 3 dígitos decimales en base 10.

- Cada clave se descompone en sus 3 dígitos: d₁ d₂ d₃
- Se usará un algoritmo LSD, de modo que, primero se clasifica por unidades (d₃), luego por decenas (d₂) y finalmente por centenas (d₁).
- Se necesitan 10 filas auxiliares.

ALGORITMO RADIX LST para claves de 3 dígitos (d₁d₂d₃):

- P0. Leer el vector
- P1. Clasificar según las unidades: d₃
- P2. Concatenar las filas en orden
- P3. Clasificar según las decenas: d₂
- P4. Concatenar las filas en orden
- P5. Clasificar según las centenas: d₁
- P6. Escribir el vector ordenado
- P7. FIN

1º. Clasificar según las unidades (dígito menos significativo)

А Г	000	000	400	00-	404	0.40	004	4 7 4	004	004
A=[03 <mark>6</mark>	909	100	025	10 1	049	064	4/1	381	004
	55 <mark>5</mark>	18 <mark>6</mark>	100	038	67 1	001	010	999	899	0021
		. • •	. • •	- - - -	• • •	• •	• . •	- - - -		

No. fila	frente				final
Fila 0	100	100	010		
Fila 1	10 <mark>1</mark>	47 <mark>1</mark>	38 <mark>1</mark>	67 <mark>1</mark>	001
Fila 2	002				
Fila 3					
Fila 4	064	004			
Fila 5	025	55 <mark>5</mark>			
Fila 6	036	186			
Fila 7					
Fila 8	038				
Fila 9	909	049	999	899	

Concatenar las filas en orden:

A=[100	100	010	101	471	381	671	001	002	064
_	004	025	555	036	186	038	909	049	999	899]

2º. Clasificar según las decenas

A=[100	100	010	1 <mark>0</mark> 1	4 <mark>7</mark> 1	3 <mark>8</mark> 1	67 1	001	002	064
_	004	0 <mark>2</mark> 5	5 <mark>5</mark> 5	036	1 <mark>8</mark> 6	038	909	049	999	8 <mark>9</mark> 9]

<u>, </u>							
Fila 0	100	100	101	001	002	004	909
Fila 1	010						
Fila 2	025						
Fila 3	036	038					
Fila 4	049						
Fila 5	5 <mark>5</mark> 5						
Fila 6	064						
Fila 7	47 1	6 7 1					
Fila 8	3 <mark>8</mark> 1	1 <mark>8</mark> 6					
Fila 9	999	899					

Concatenar las filas en orden:

A=[100	100	101	001	002	004	909	010	025	036
_	038	049	555	064	471	671	381	186	999	899]

3°. Clasificar según las centenas

A=[100	100	1 01	001	002	004	909	0 10	025	036
_	038	049	5 55	<mark>0</mark> 64	4 71	6 71	3 81	1 86	999	<mark>8</mark> 99]

Fila 0	001	002	004	0 10	025	036	038	049	064
Fila 1	100	100	1 01	1 86					
Fila 2									
Fila 3	3 81								
Fila 4	4 71								
Fila 5	5 55								
Fila 6	6 71								
Fila 7									
Fila 8	<mark>8</mark> 99								
Fila 9	909	999							

Concatenar las filas en orden:

A=[001	002	004	010	025	036	038	049	064	100
_	100	101	186	381	471	555	671	899	909	999]

Algoritmo Radix

Algoritmo Radix (A,n,M)

Entrada:

A: arreglo de (1.. MAX) elementos de enteros en base 10

n: nro. entero de datos a ordenar

M: cantidad de dígitos de la clave

Salida:

A: arreglo ordenado.

Auxiliar:

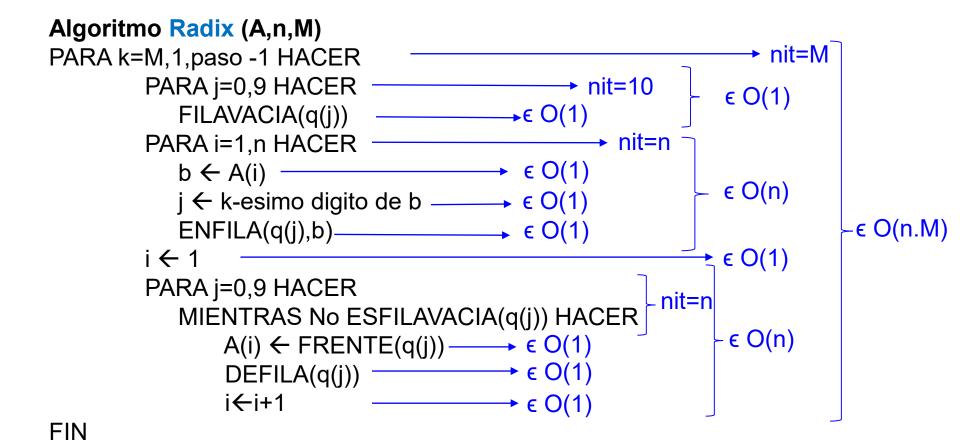
b, exp, i, j, k € enteros, q(10) € arreglo de filas de enteros (base 10)

Algoritmo Radix

```
Algoritmo Radix (A,n,M)
PARA k=M,1,paso -1 HACER
        PARA j=0,9 HACER
                 FILAVACIA(q(j))
        PARA i=1,n HACER
                 b \leftarrow A(i)
                 j ← k-esimo digito de b
                 ENFILA(q(j),b)
        i ← 1
        PARA j=0,9 HACER
                 MIENTRAS No ESFILAVACIA(q(j)) HACER
                          A(i) \leftarrow FRENTE(q(j))
                          DEFILA(q(j))
                          i←i+1
```

FIN

Costo Algoritmo Radix



36

```
A_0 = (ASORTINGEXAMPLE)
Elegir un código binario para cada letra:
A:00001, B:00010, C:00011, D:00100, E:00101 ...
                                                         01001
K=5
                                                         00111
                                                         00101
                                                         11000
Filas:
                                                         00001
q(0)=
                                                         10000
q(1)=
                                                         00101
```

 A_0 =(ASORTINGEXAMPLE)

K=5

Llevar las claves a las Filas:

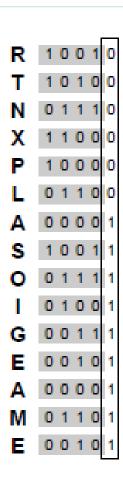
q(0): RTNXPL

q(1): ASOIGEAME

Rearmar el Arreglo:

 A_1 =(RTNXPLASOIGEAME)

				•	Ţ	
Α	0	0	0	0	1	
S	1	0	0	1	1	
0	0	1	1	1	1	
R	1	0	0	1	0	
T	1	0	1	0	0	
	0	1	0	0	1	
Ν	0	1	1	1	0	
G	0	0	1	1	1	
E	0	0	1	0	1	
X	1	1	0	0	0	
Α	0	0	0	0	1	
M	0	1	1	0	1	
P	1	0	0	0	0	
L	0	1	1	0	0	
E	0	0	1	0	1	



 $A_1 = (RTNXPLASOIGEAME)$

K=4

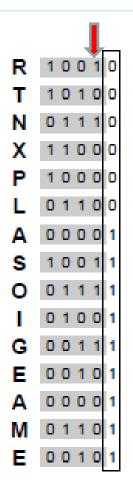
Llevar las claves a las Filas:

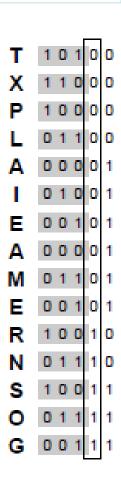
q(0): TXPLAIEAME

q(1): RNSOG

Rearmar el Arreglo:

 A_2 =(TXPLAIEAMERNSOG)





 A_2 =(TXPLAIEAMERNSOG)

K=3

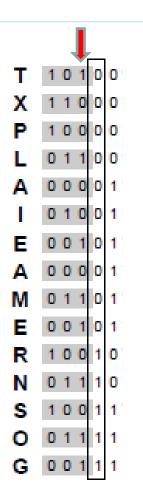
Llevar las claves a las Filas:

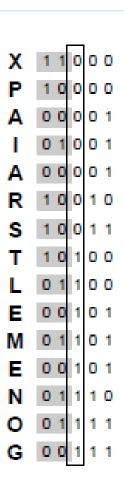
q(0): XPAIARS

q(1): TLEMENOG

Rearmar el Arreglo:

 A_3 =(XPAIARSTLEMENOG)





 A_3 =(XPAIARSTLEMENOG)

K=2

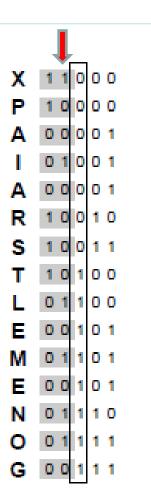
Llevar las claves a las Filas:

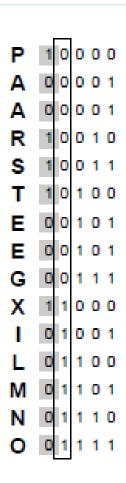
q(0): PAARSTEEG

q(1): XILMNO

Rearmar el Arreglo:

 A_4 =(PAARSTEEGXILMNO)





 A_{Δ} =(PAARSTEEGXILMNO)

K=1

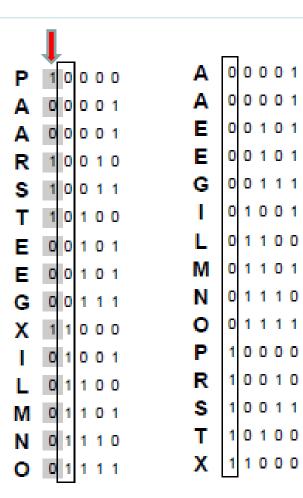
Llevar las claves a las Filas:

q(0): AAEEGILMNO

q(1): PRSTX

Rearmar el Arreglo:

 $A_5 = (AAEEGILMNOPRSTX)$



												l	
Α	00001	R	1 0 0 1 0	Т	101001	X	1 1 0 0 0	P	10000	Α	0	0 (0 0 1
S	10011	T	1 0 1 0 0	X	11000	P	10000	Α	0 0 0 0 1	Α	0	0 (0 0 1
0	01111	N	0 1 1 1 0	Р	10000	A	0 0 0 0 1	Α	0 0 0 0 1	E	0	0	1 0 1
R	10010	X	1 1 0 0 0	L	01100	/ I	0 1 0 0 1	R	10010	E	0	0	101
T	10100	P	1 0 0 0 0	Α	000011/	/ A	0 0 0 0 1	S	10011	G	0	0	1 1 1
I	01001	L	0 1 1 0 0	- 1	01001/	/ R	10010	T	10100	-	0	1 (001
Ν	0 1 1 1 0	Α	00001	Ε	00101	S	10011	Ε	0 0 1 0 1	L	0	1	100
G	00111	S	1 0 0 1 1	Α	00001/	T	10100	Ε	0 0 1 0 1	M	0	1	101
E	00101	0	0 1 1 1 1	M	01101	L	0 1 1 0 0	G	0 0 1 1 1	Ν	0	1	1 1 0
X	11000	I	0 1 0 0 1	E	00101	E	0 0 1 0 1	X	1 1 0 0 0	0	0	1	1 1 1
Α	00001	G	0 0 1 1 1	R	10010	M	0 1 1 0 1	- 1	0 1 0 0 1	P	1	0 (000
M	01101	E	00101	N	01110	`E	0 0 1 0 1	L	0 1 1 0 0	R	1	0 (0 1 0
P	10000	Α	0 0 0 0 1	S	10011	N,	0 1 1 1 0	М	0 1 1 0 1	S	1	0 (1 1
L	01100	M	0 1 1 0 1	0	0 1 1 1 1	0	0 1 1 1 1	N	0 1 1 1 0	Т	1	0	100
Ε	00101	Ε	0 0 1 0 1	G	0 0 1 1	G	0 0 1 1 1	0	0 1 1 1 1	X	1	1 (000

Algoritmo Radix

Tiempo de Ejecución:

Para ordenar un vector de *n* datos de *M* dígitos cada uno:

 $T(n) \in O(n \cdot M)$.

Memoria: Si los datos están en base **b**, se usa espacio extra para implementar las **b** filas.

Sensibilidad: nada sensible.

Estabilidad: si es estable

Otros

- Puede usarse para listas enlazadas
- Puede diseñarse una versión recursiva.
- Puede paralelizarse en su versión recursiva.