

Algoritmos e Estruturas de Dados II

Aula 21: Ordenação Externa - Parte 2

Prof. Jurandy G. Almeida Jr.

Universidade Federal de São Paulo
Departamento de Ciência e Tecnologia

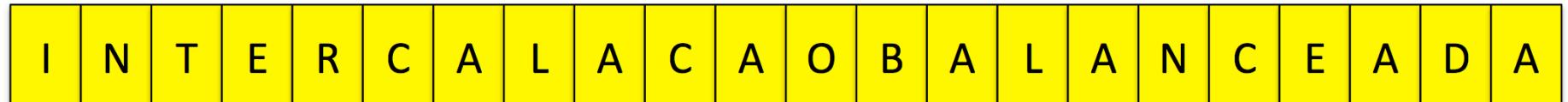
29 de Junho de 2016

ORDENAÇÃO EXTERNA
INTERCALAÇÃO POLIFÁSICA

- Problemas com a Intercalação Balanceada:
 - Necessita de um grande número de fitas, fazendo várias leituras e escritas entre as fitas envolvidas
 - Para uma intercalação balanceada de f -caminhos, geralmente são necessárias $2f$ fitas
 - Alternativamente, pode-se copiar o arquivo de uma única fita de saída para f fitas de entrada, reduzindo o número de fitas para $f + 1$
 - Há um custo de uma cópia adicional do arquivo
- Solução: **Intercalação Polifásica**

- Processo de funcionamento da intercalação polifásica:
 - Os blocos ordenados são distribuídos de forma desigual entre as fitas disponíveis
 - Uma fita é deixada livre
 - Em seguida, a intercalação de blocos ordenados é executada até que uma das fitas de entrada se esvazie
 - A fita vazia torna-se a próxima fita de saída

Intercalação Polifásica



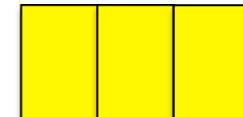
Fita 1

Fita 2

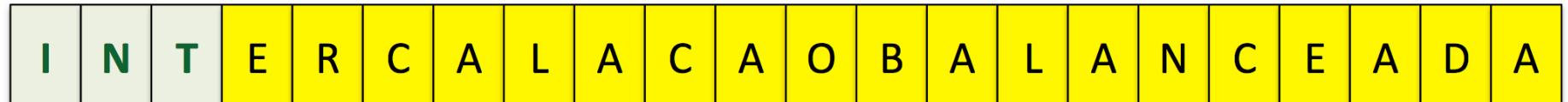
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1

Fita 2

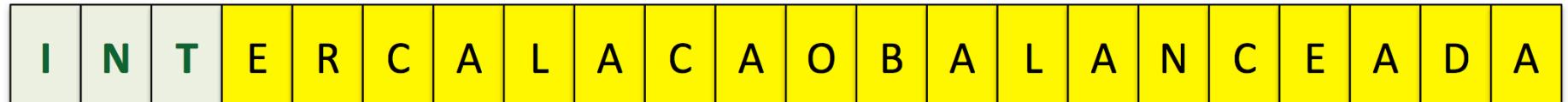
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1

Fita 2

Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1

Fita 2

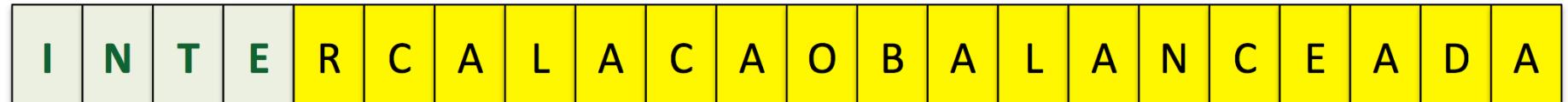
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1



Fita 2

Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1



Fita 2

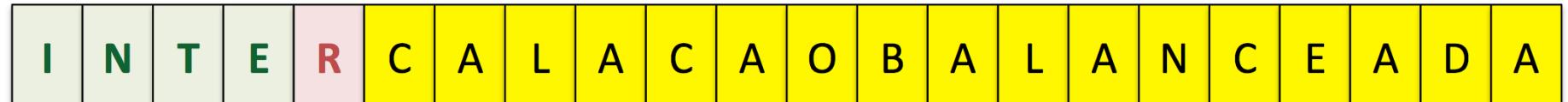
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1



Fita 2

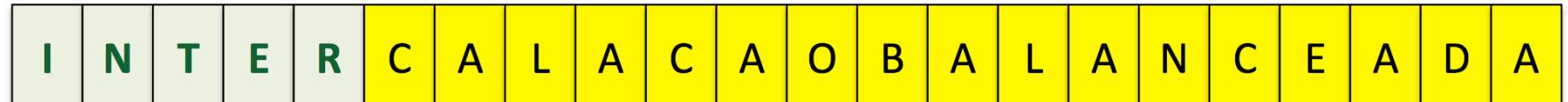
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1  

Fita 2

Fita 3

Memória Principal

1 2 3



R E T

Intercalação Polifásica



Fita 1

I	N
---	---

Fita 2

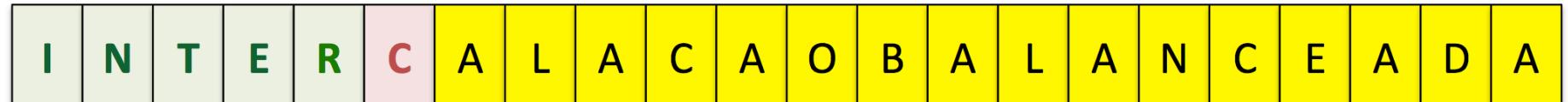
Fita 3

Memória Principal

1 2 3

R	E	T
---	---	---

Intercalação Polifásica



Fita 1

I	N
---	---

Fita 2

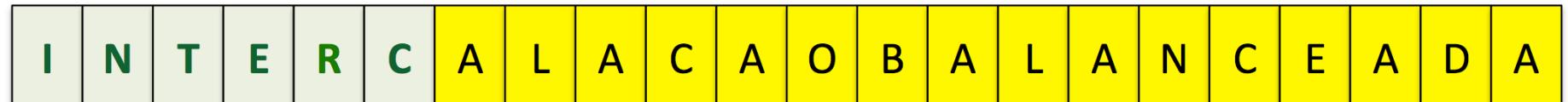
Fita 3

Memória Principal

1 2 3

R	E	T
---	---	---

Intercalação Polifásica



Fita 1

I	N	R
---	---	---

Fita 2

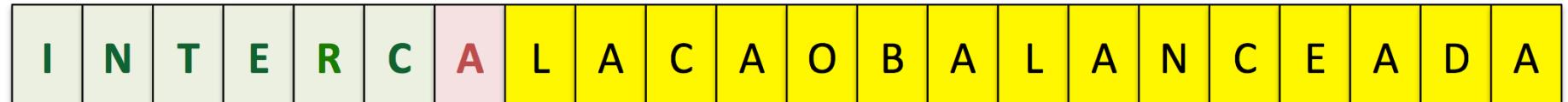
Fita 3

Memória Principal

1 2 3

T	E	C
---	---	---

Intercalação Polifásica



Fita 1

I	N	R
---	---	---

Fita 2

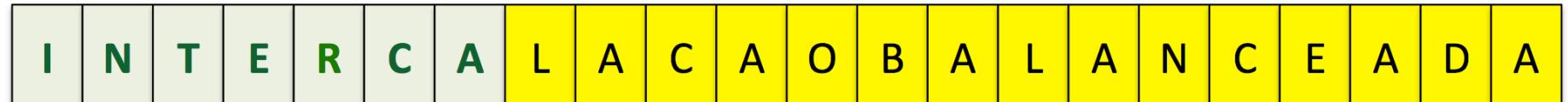
Fita 3

Memória Principal

1 2 3

T	E	C
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

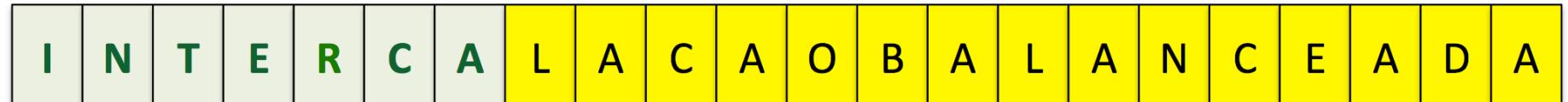
Fita 3

Memória Principal

1 2 3

A	C	E
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

Fita 3

Memória Principal

1 2 3

A	C	E
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

Fita 3

Memória Principal

1 2 3

A	C	E
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A

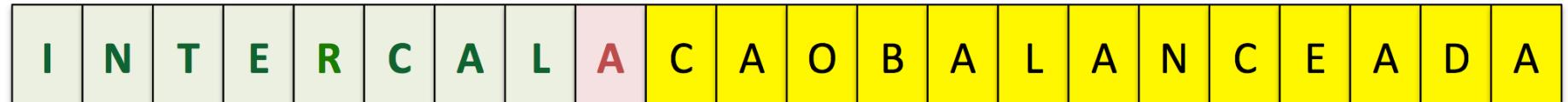
Fita 3

Memória Principal

1 2 3

C	L	E
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A

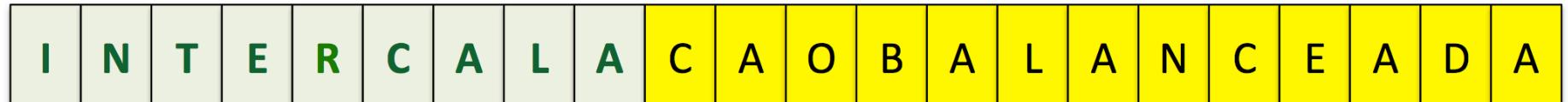
Fita 3

Memória Principal

1 2 3

C	L	E
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A	C
---	---

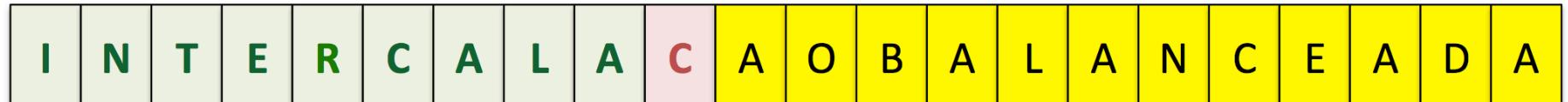
Fita 3

Memória Principal

1 2 3

E	L	A
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A	C
---	---

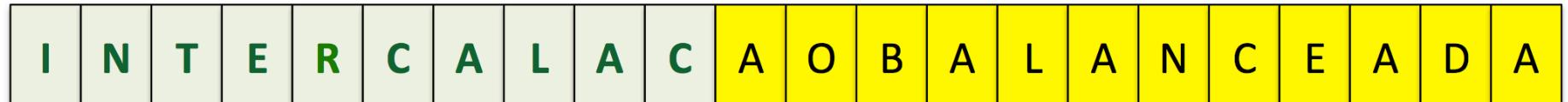
Fita 3

Memória Principal

1 2 3

E	L	A
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A	C	E
---	---	---

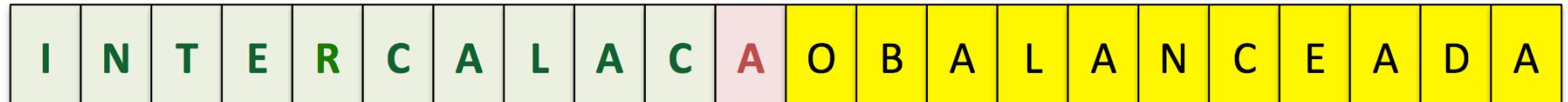
Fita 3

Memória Principal

1 2 3

L	C	A
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A	C	E
---	---	---

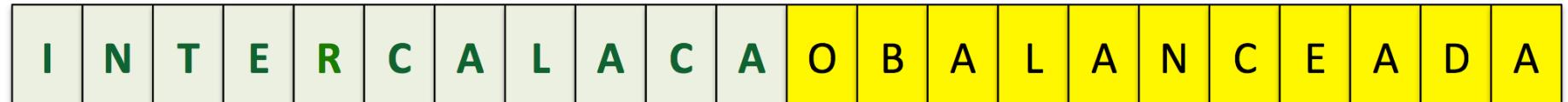
Fita 3

Memória Principal

1 2 3

L	C	A
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A	C	E	L
---	---	---	---

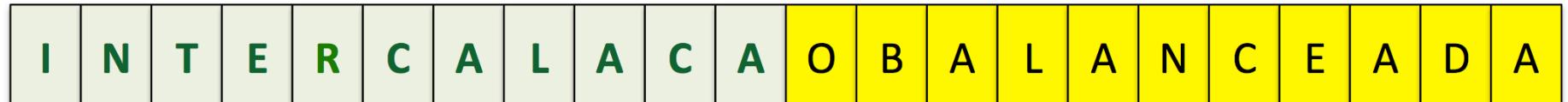
Fita 3

Memória Principal

1 2 3

A	C	A
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A	C	E	L
---	---	---	---

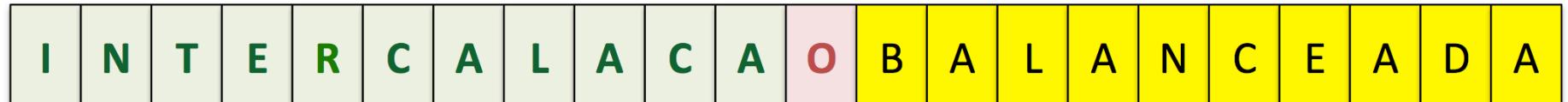
Fita 3

Memória Principal

1 2 3

A	C	A
---	---	---

Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

Fita 2

A	C	E	L
---	---	---	---

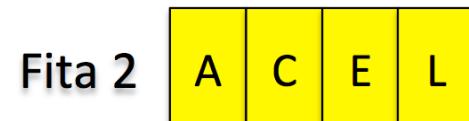
Fita 3

Memória Principal

1 2 3

A	C	A
---	---	---

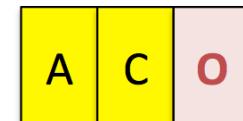
Intercalação Polifásica



Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

 A

Fita 2

A	C	E	L
---	---	---	---

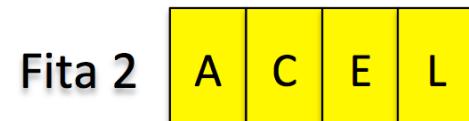
Fita 3

Memória Principal

1 2 3

A	C	O
---	---	---

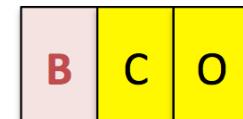
Intercalação Polifásica



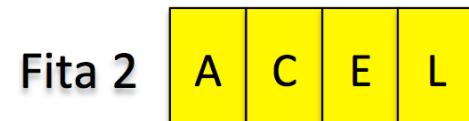
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



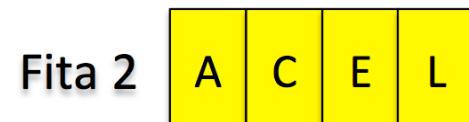
Fita 3

Memória Principal

1 2 3



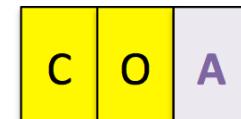
Intercalação Polifásica



Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



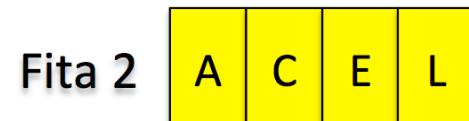
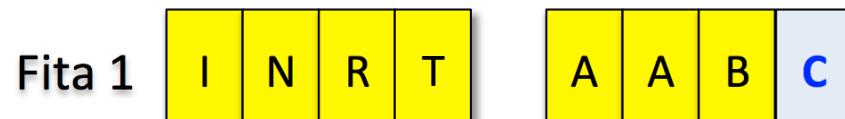
Fita 3

Memória Principal

1 2 3



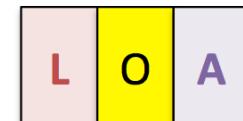
Intercalação Polifásica



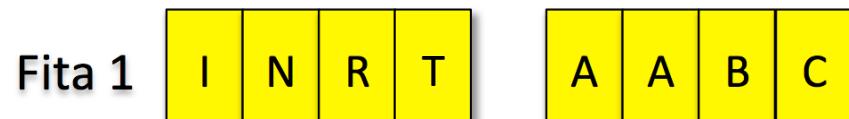
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



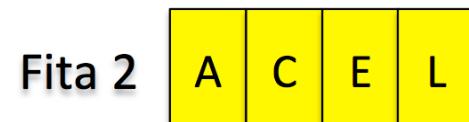
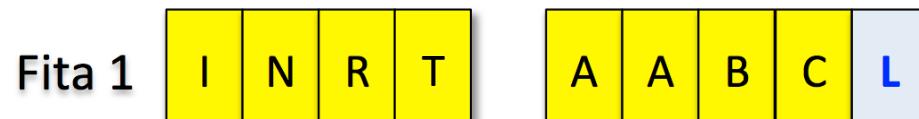
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



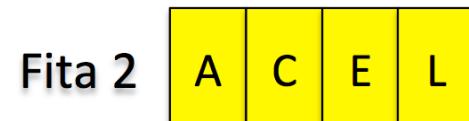
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



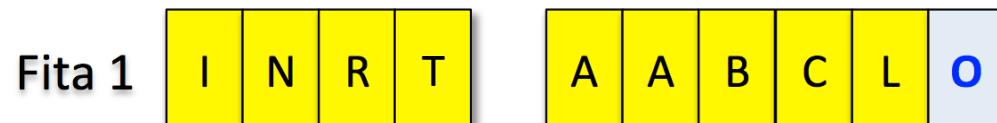
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



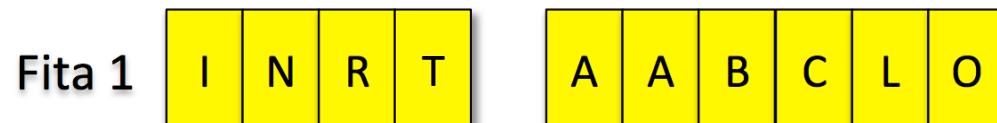
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 1

I	N	R	T
---	---	---	---

A	A	B	C	L	O
---	---	---	---	---	---

Fita 2

A	C	E	L
---	---	---	---

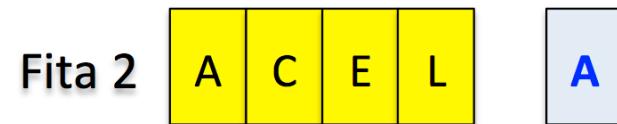
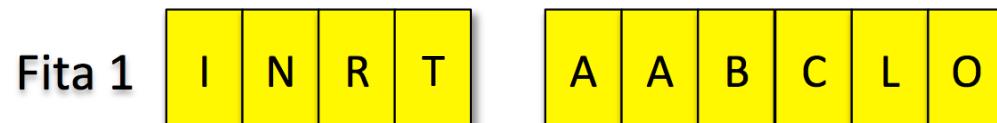
Fita 3

Memória Principal

1 2 3

A	N	A
---	---	---

Intercalação Polifásica



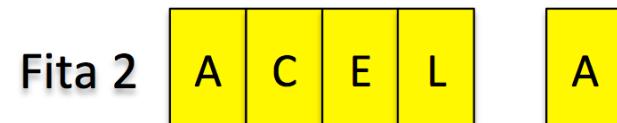
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



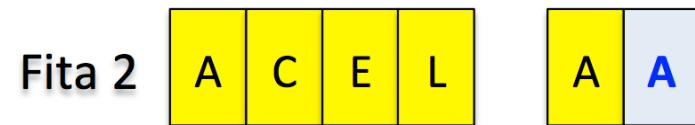
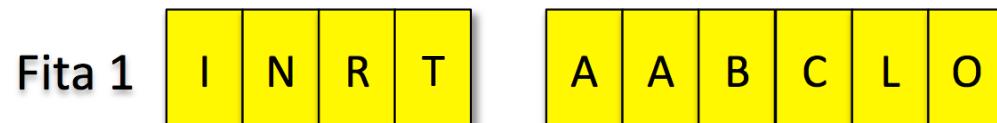
Fita 3

Memória Principal

1 2 3



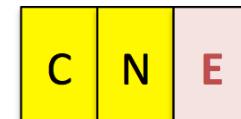
Intercalação Polifásica



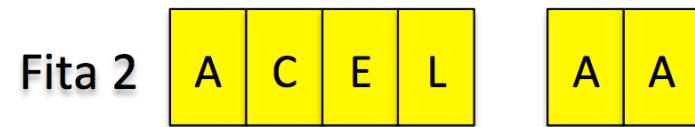
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



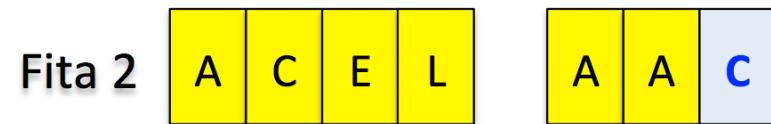
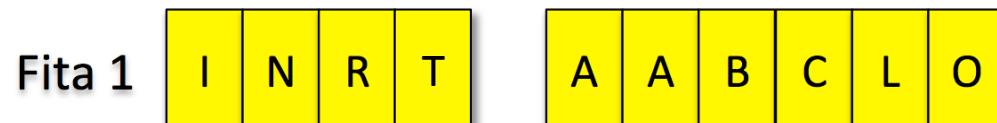
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



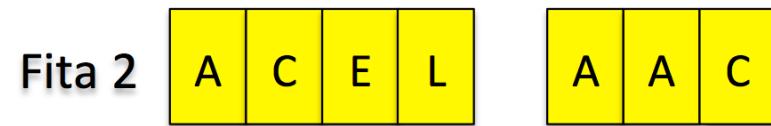
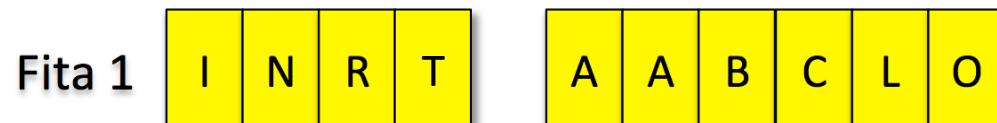
Fita 3

Memória Principal

1 2 3



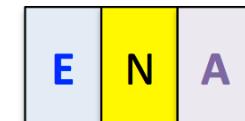
Intercalação Polifásica



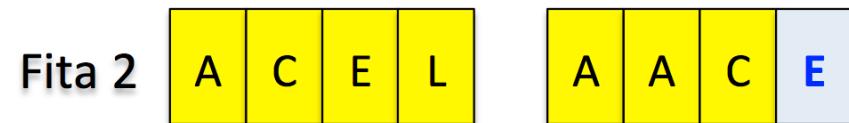
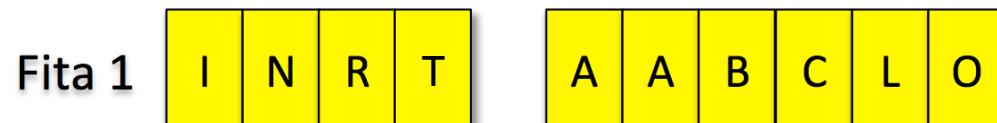
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



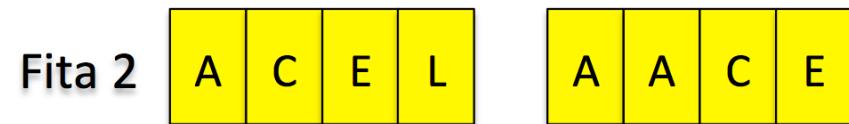
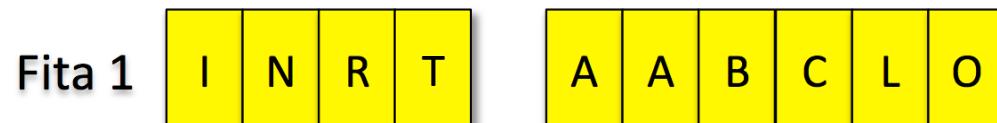
Fita 3

Memória Principal

1 2 3



Intercalação Polifásica



Fita 3

Memória Principal

1 2 3



Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

Fita 1 I N R T A A B C L O

Fita 2 A C E L A A C E N

Fita 3

Memória Principal

1 2 3

A D A

Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

Fita 1 I N R T A A B C L O

Fita 2 A C E L A A C E N

Fita 3

Memória Principal

1 2 3

A D A

Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

Fita 1 I N R T A A B C L O

Fita 2 A C E L A A C E N

Fita 3

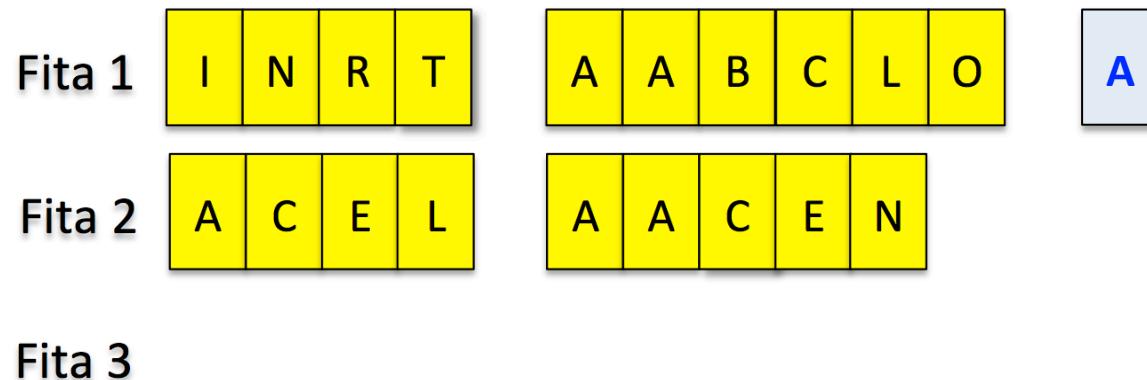
Memória Principal

1 2 3

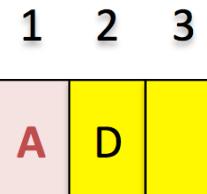
A	D	A
---	---	---

Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A



Memória Principal



Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

Fita 1 I N R T A A B C L O A

Fita 2 A C E L A A C E N

Fita 3

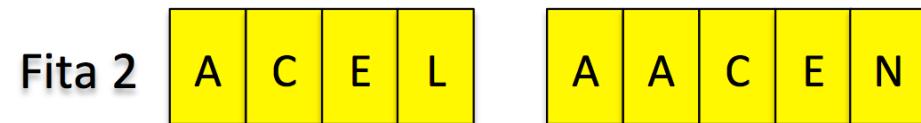
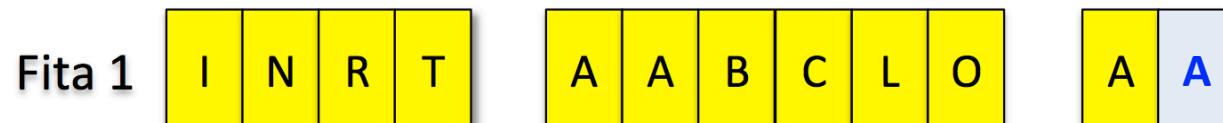
Memória Principal

1 2 3



Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A



Fita 3

Memória Principal

1 2 3



Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

Fita 1 I N R T A A B C L O A A

Fita 2 A C E L A A C E N

Fita 3

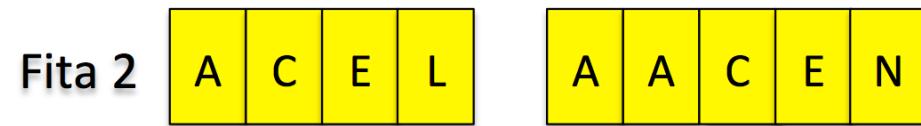
Memória Principal

1 2 3



Intercalação Polifásica

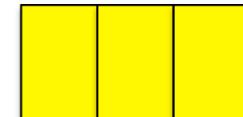
I N T E R C A L A C A O B A L A N C E A D A



Fita 3

Memória Principal

1 2 3



Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

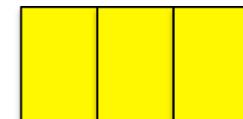
Fita 1 I N R T A A B C L O A A D

Fita 2 A C E L A A C E N

Fita 3

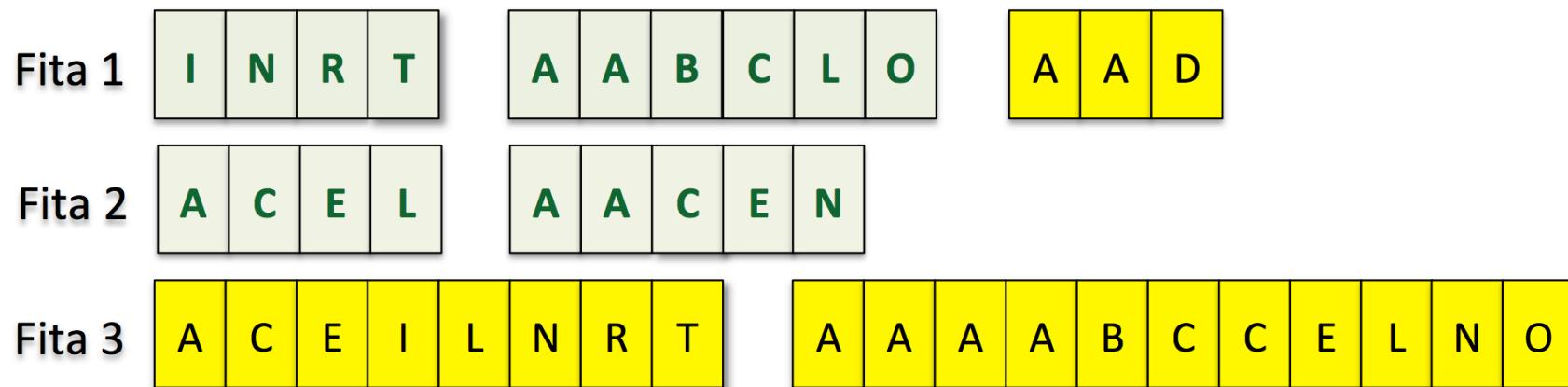
Memória Principal

1 2 3



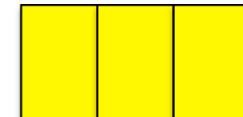
Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A



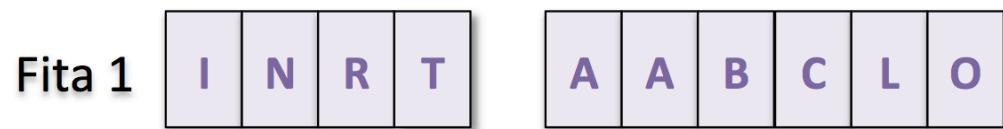
Memória Principal

1 2 3

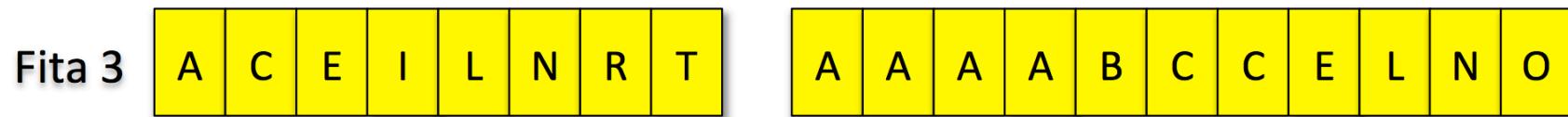


Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

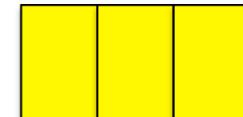


Fita 2



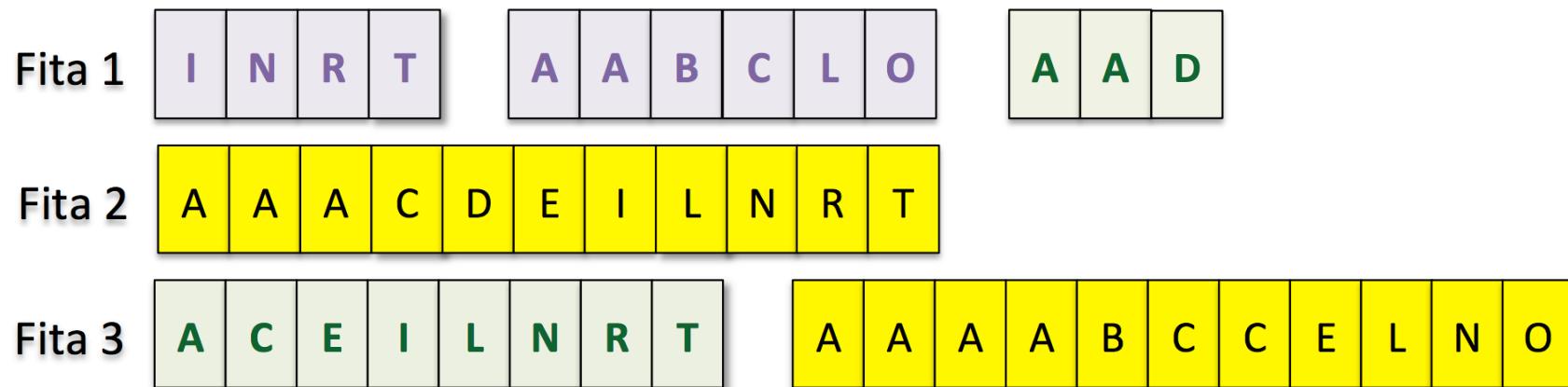
Memória Principal

1 2 3



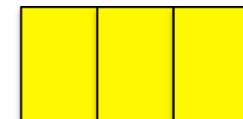
Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

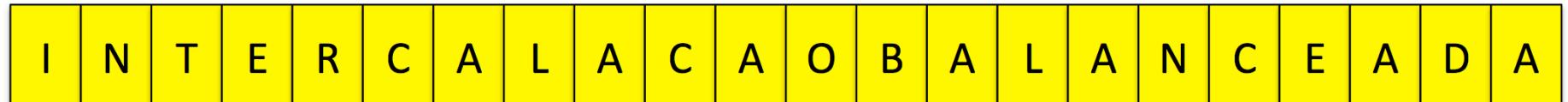


Memória Principal

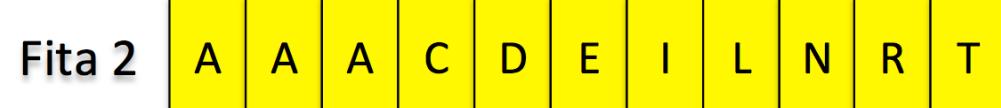
1 2 3



Intercalação Polifásica

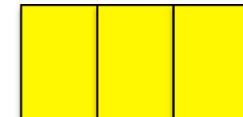


Fita 1



Memória Principal

1 2 3



Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

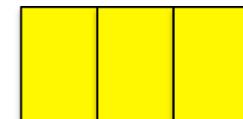
Fita 1 A A A A A A A B C C C D E E I L L N N O R T

Fita 2 A A A C D E I L N R T

Fita 3 A C E I L N R T A A A A B C C E L N O

Memória Principal

1 2 3



Intercalação Polifásica

I N T E R C A L A C A O B A L A N C E A D A

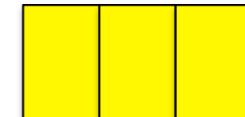
Fita 1 A A A A A A A B C C C D E E I L L N N O R T

Fita 2

Fita 3

Memória Principal

1 2 3



- A intercalação é realizada muitas vezes
- As fases não envolvem todos os blocos
- Nenhuma cópia direta entre fitas é realizada
- A implementação da intercalação polifásica é simples
- A parte mais delicada está na distribuição inicial dos blocos ordenados entre as fitas

- A análise da intercalação polifásica é complicada
- O que se sabe é que ela é ligeiramente melhor do que a intercalação balanceada para valores pequenos de f
- Para valores de $f > 8$, a intercalação balanceada de vários caminhos pode ser mais rápida

ORDENAÇÃO EXTERNA
QUICKSORT EXTERNO

QuickSort Externo

- *QuickSort* externo foi proposto por Monard em 1980
- O algoritmo utiliza o paradigma de divisão e conquista
- O algoritmo realiza a ordenação *in-place* de um arquivo $\mathbf{A} = \{R_1, \dots, R_n\}$ de n registros
 - Os registros encontram-se armazenados consecutivamente em memória secundária de acesso randômico
- O algoritmo utiliza somente $\log n$ unidades de memória interna, não sendo necessária alguma memória externa adicional

QuickSort Externo

- Para ordenar o arquivo $\mathbf{A} = \{R_1, \dots, R_n\}$, o algoritmo:
 - particiona A da seguinte forma:

$$\{R_1, \dots, R_i\} \leq R_{i+1} \leq R_{i+2} \leq \dots \leq R_{j-2} \leq R_{j-1} \leq \{R_j, \dots, R_n\}$$
 - chama recursivamente o algoritmo em cada um dos subarquivos gerados:

$$\mathbf{A}_1 = \{R_1, \dots, R_i\} \text{ e } \mathbf{A}_2 = \{R_j, \dots, R_n\}$$
- Os registros ordenados $\{R_{i+1}, \dots, R_{j-1}\}$ correspondem ao pivô do algoritmo, encontrando-se na memória interna durante a execução do mesmo
 - Os subarquivos \mathbf{A}_1 e \mathbf{A}_2 contêm os registros menores que R_{i+1} e maiores que R_{j-1} , respectivamente

QuickSort Externo

- Para a partição do arquivo, é utilizada uma área de memória interna para o armazenamento do pivô
 - Tamanho da área: $m = j - i - 1$, com $m \geq 3$
- Nas chamadas recursivas, deve-se considerar que:
 - deve ser ordenado, inicialmente, o subarquivo de menor tamanho
 - subarquivos vazios ou com um único registro são ignorados
 - caso o arquivo de entrada **A** possua no máximo m registros, ele é ordenado em um único passo

QuickSort Externo

- Considerações sobre o funcionamento de uma partição:
 - Os valores das chaves de R_i e R_j são denominados limite inferior (L_{inf}) e limite superior (L_{sup}) da partição
 - Tais limites são inicializados com os valores $-\infty$ e $+\infty$
 - A leitura de A é controlada por apontadores de leitura inferior (L_i) e superior (L_s).
 - A cada leitura no extremo inferior, L_i é incrementado; a cada leitura no extremo superior, L_s é decrementado
 - A escrita em A é controlada por apontadores de escrita inferior (E_i) e superior (E_s)
 - A cada escrita no extremo inferior, E_i é incrementado; a cada escrita no extremo superior, E_s é decrementado

QuickSort Externo

- Os primeiros $m - 1$ registros são lidos, alternativamente, dos extremos de \mathbf{A} e armazenados na área de memória interna
- Ao ler o m -ésimo registro, cuja chave é C :
 - C é comparada com L_{sup} e, sendo maior, j recebe E_s e o registro é escrito em A_2
 - caso contrário, C é comparada com L_{inf} e, sendo menor, i recebe E_i e o registro é escrito em A_1
 - Caso contrário ($L_{inf} \leq C \leq L_{sup}$), o registro é inserido na área de memória interna

QuickSort Externo

- Para garantir que os apontadores de escrita estejam atrás dos apontadores de leitura, a ordem alternada de leitura é interrompida se ($L_i = E_i$) ou ($L_s = E_s$)
 - Nenhum registro pode ser destruído durante a ordenação *in-place*

QuickSort Externo

- Quando a área de memória enche, deve-se remover um registro da mesma, considerando os tamanhos atuais de A_1 e A_2
 - Sendo Esq e Dir a primeira e a última posição de A , os tamanhos de A_1 e A_2 são, respectivamente, ($T_1 = E_i - Esq$) e ($T_2 = Dir - E_s$)
 - Se ($T_1 < T_2$), o registro de menor chave é removido da memória, sendo escrito em E_i (A_1), e L_{inf} é atualizado com tal chave
 - Se ($T_1 \geq T_2$), o registro de maior chave é removido da memória, sendo escrito em E_s (A_2), e L_{sup} é atualizado com tal chave

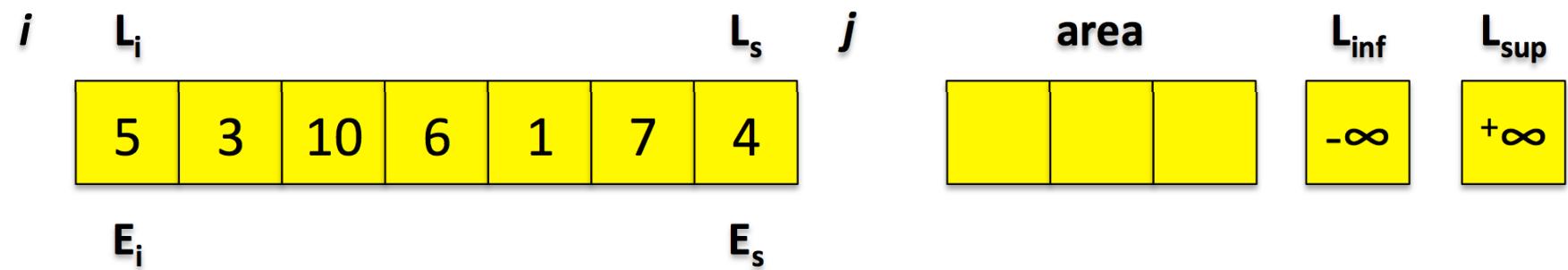
QuickSort Externo

- O objetivo é escrever o registro removido da memória no subarquivo de menor tamanho, no intuito de dividir **A** de forma uniforme e, assim, balancear a árvore gerada pelas recursões
 - Isso minimiza a quantidade de operações de leitura e escrita efetuadas pelo algoritmo

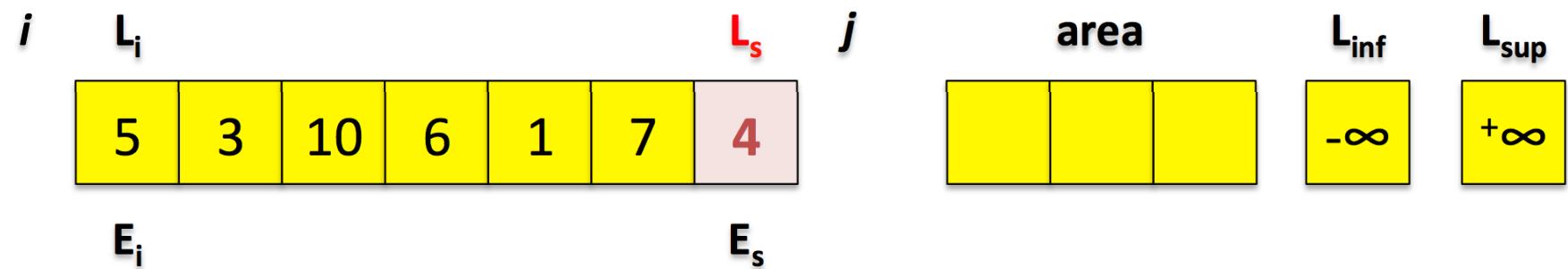
QuickSort Externo

- O processo de partição continua até que L_i e L_s se cruzem, ou seja, ($L_s < L_i$)
- Neste momento, os registros armazenados na área de memória interna devem ser copiados, já ordenados, em **A**
 - Enquanto existir registros na área de memória, o menor deles é removido e escrito na posição indicada por E_i em **A**

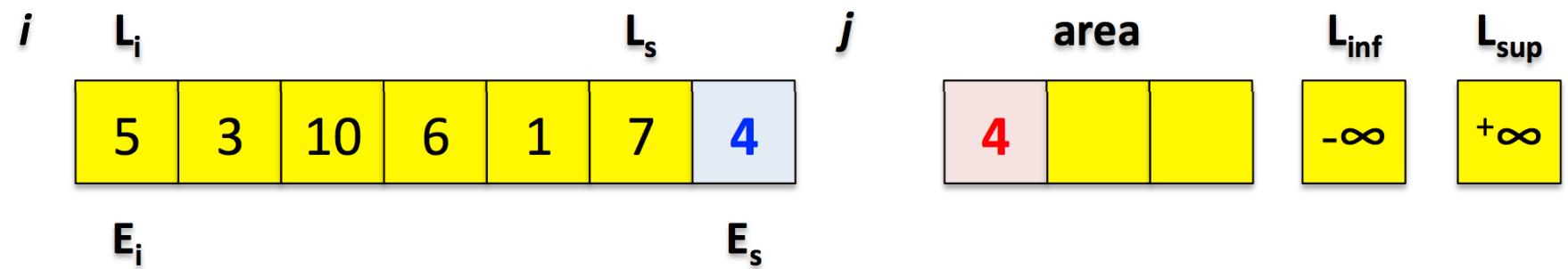
QuickSort Externo



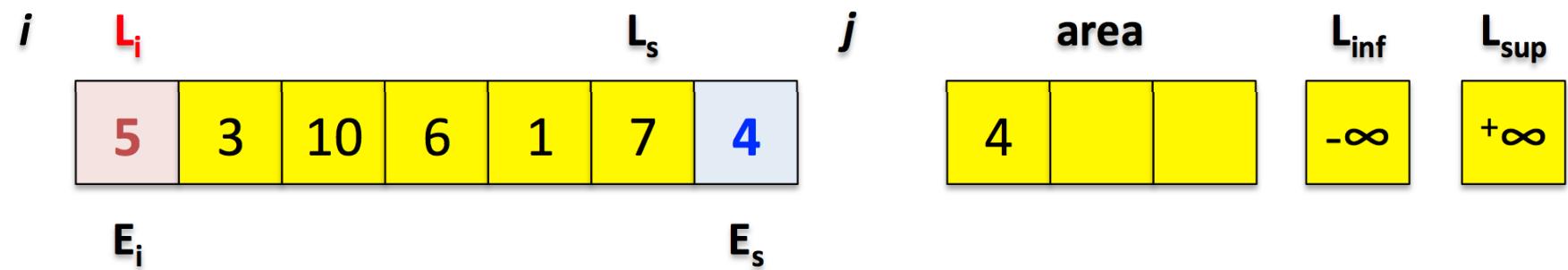
QuickSort Externo



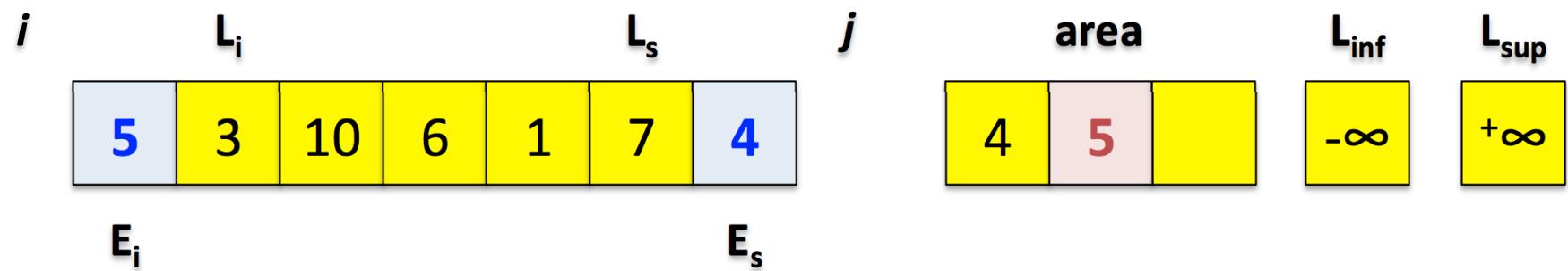
QuickSort Externo



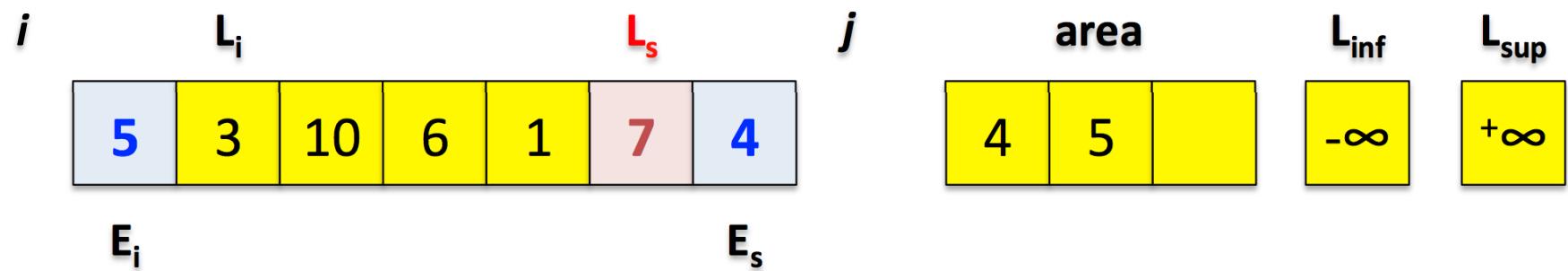
QuickSort Externo



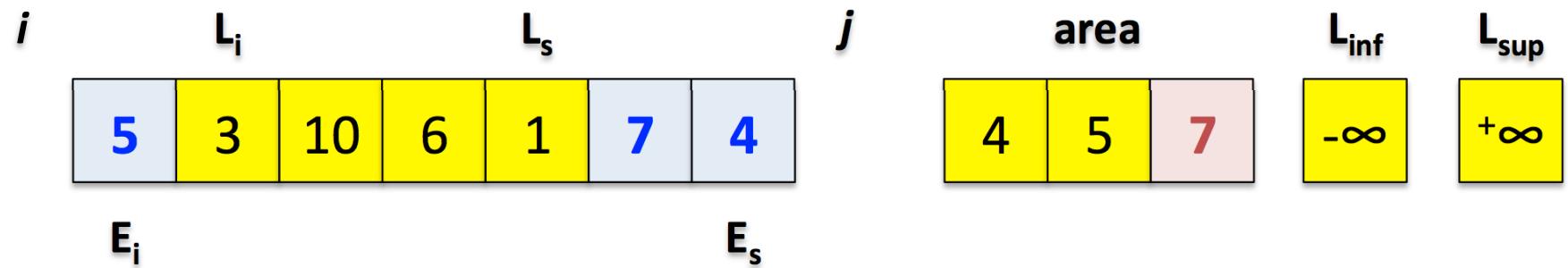
QuickSort Externo



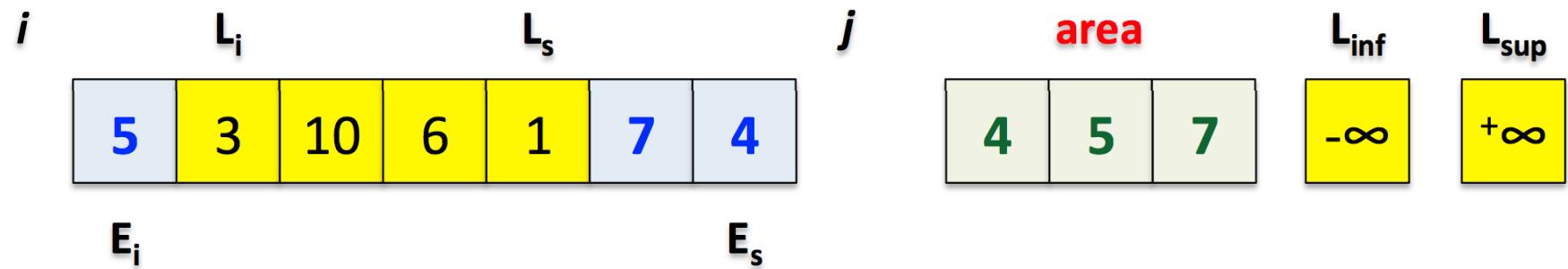
QuickSort Externo



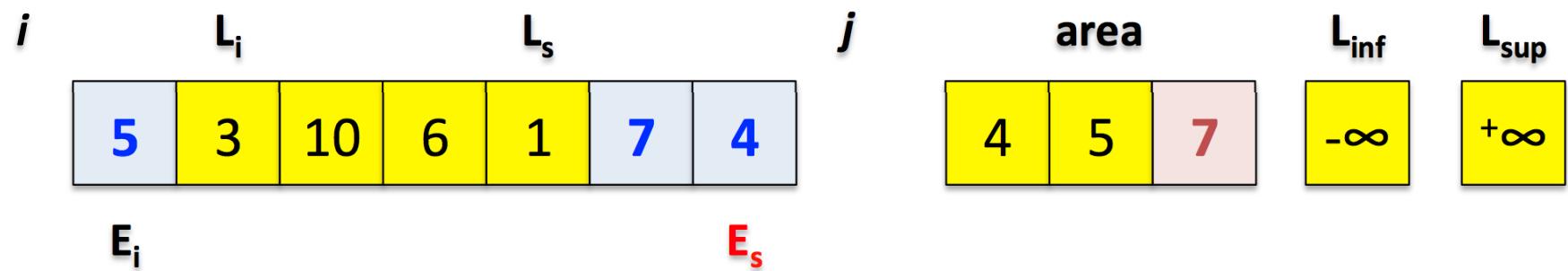
QuickSort Externo



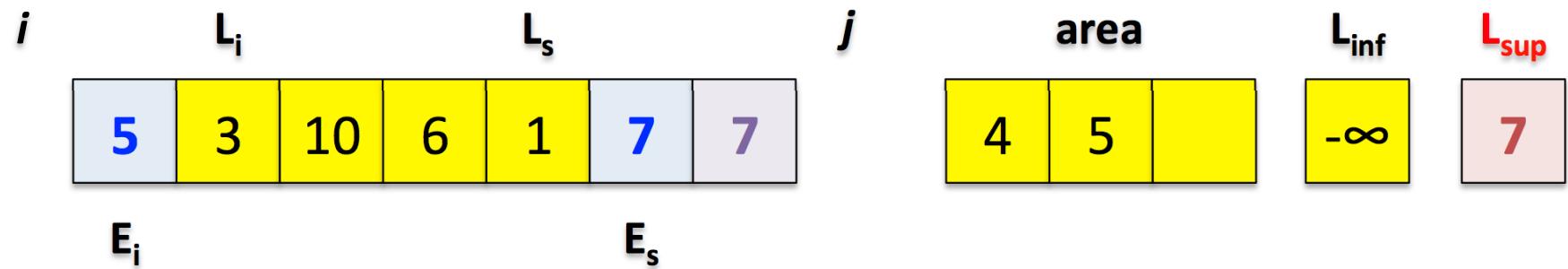
QuickSort Externo



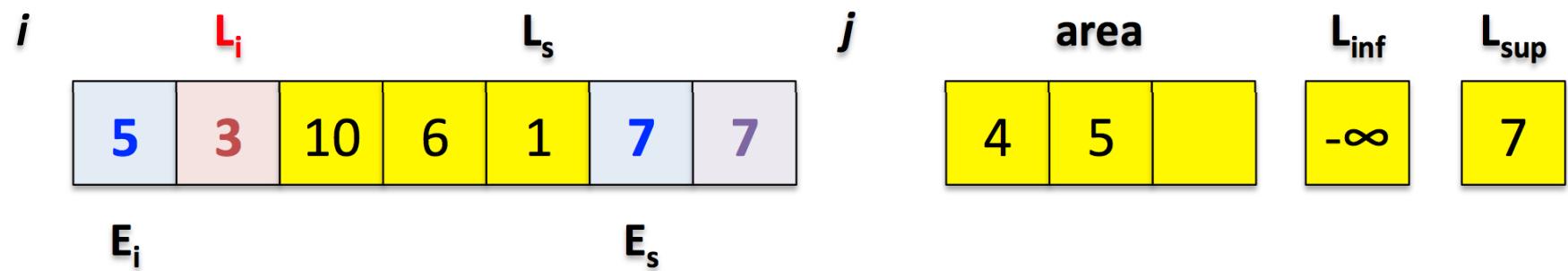
QuickSort Externo



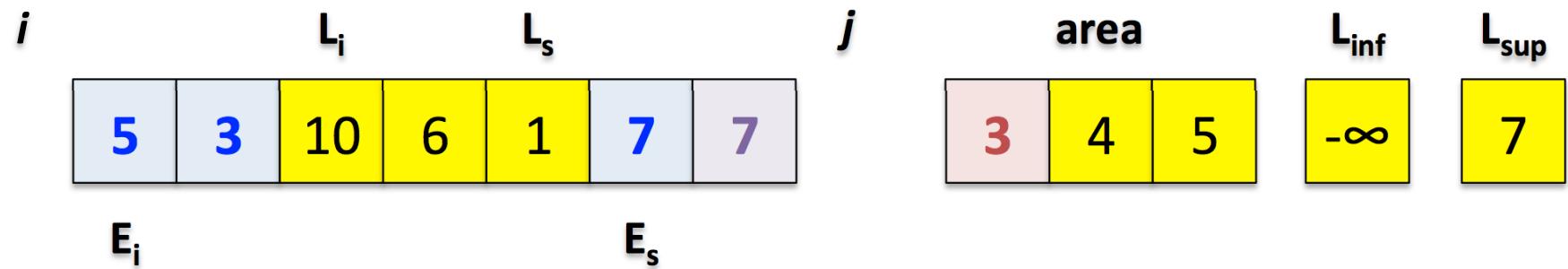
QuickSort Externo



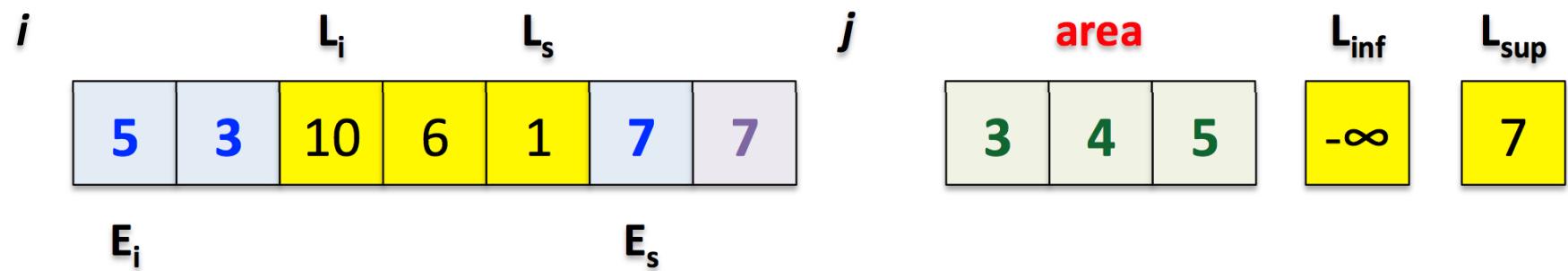
QuickSort Externo



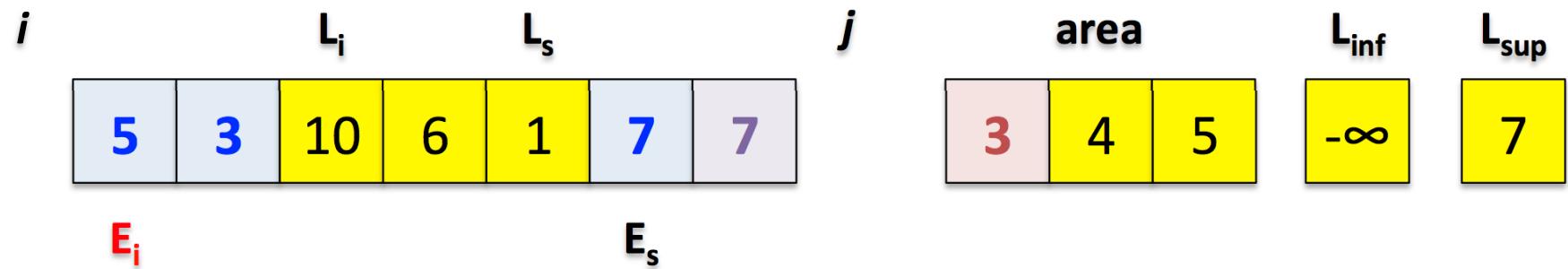
QuickSort Externo



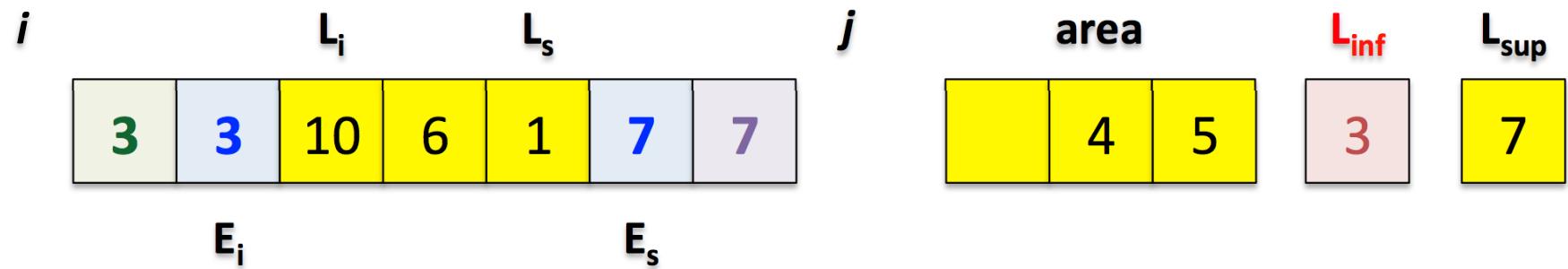
QuickSort Externo



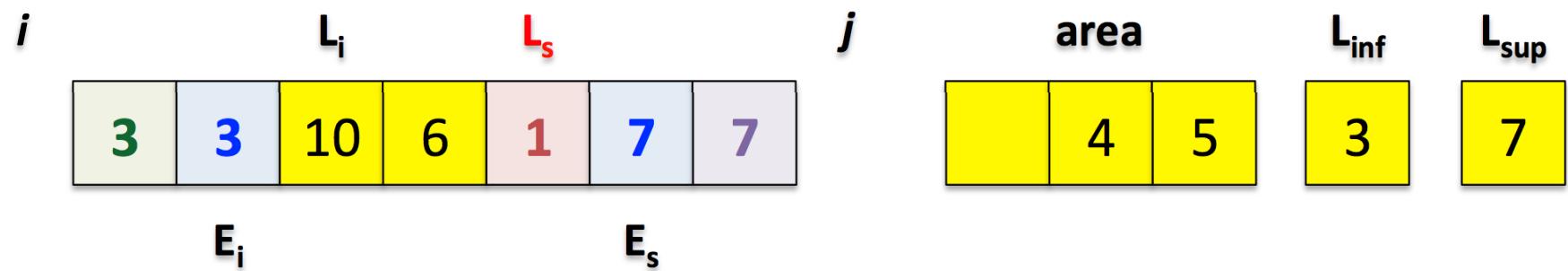
QuickSort Externo



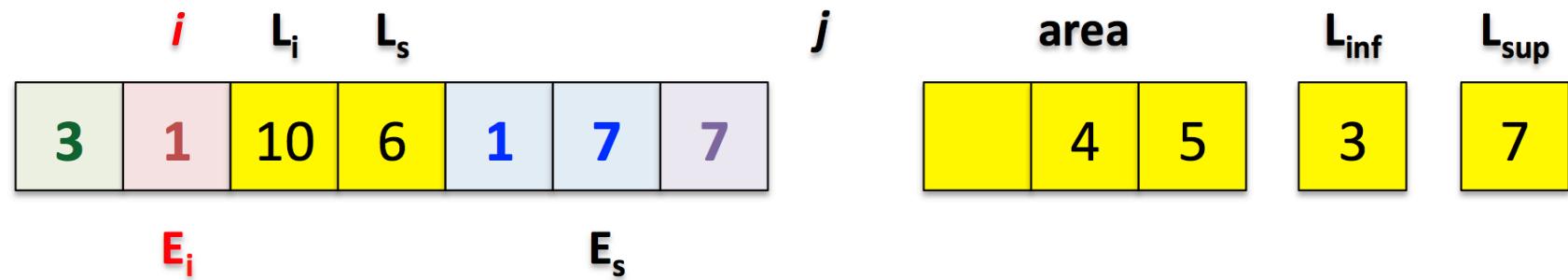
QuickSort Externo



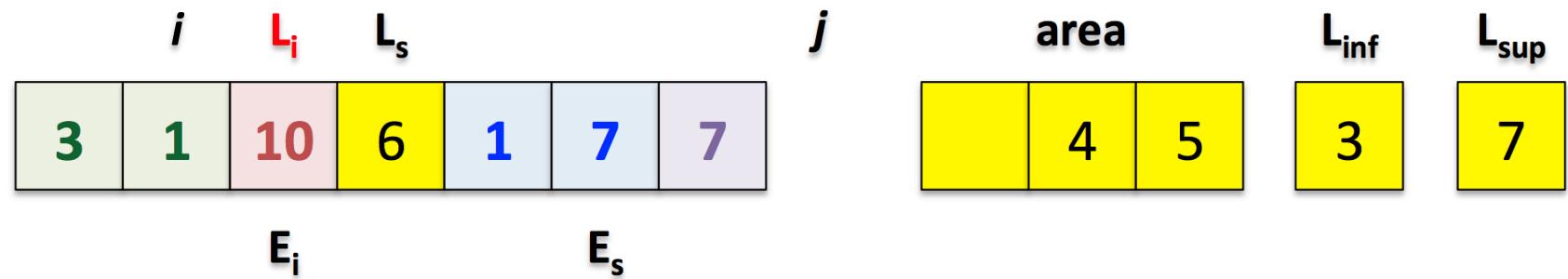
QuickSort Externo



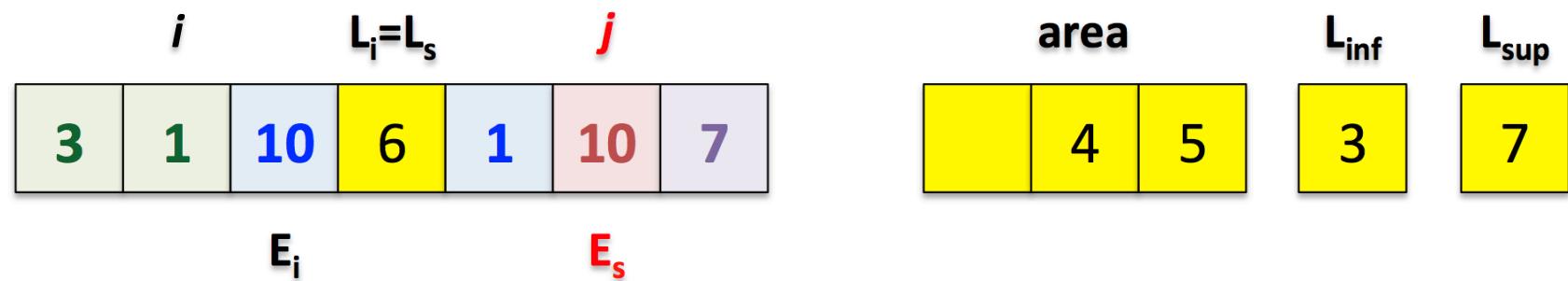
QuickSort Externo



QuickSort Externo



QuickSort Externo



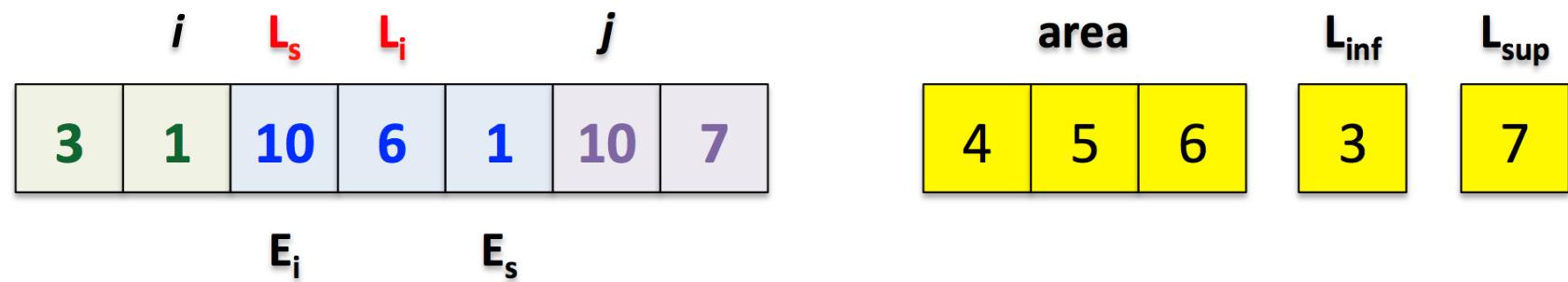
QuickSort Externo



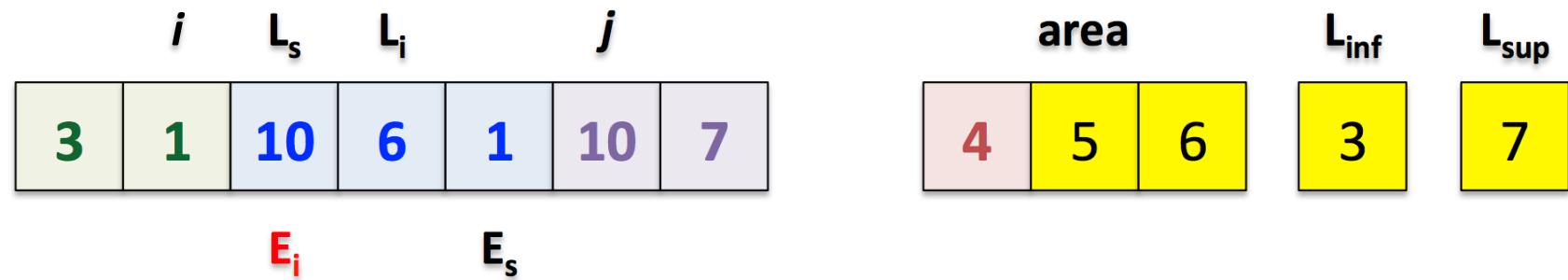
QuickSort Externo



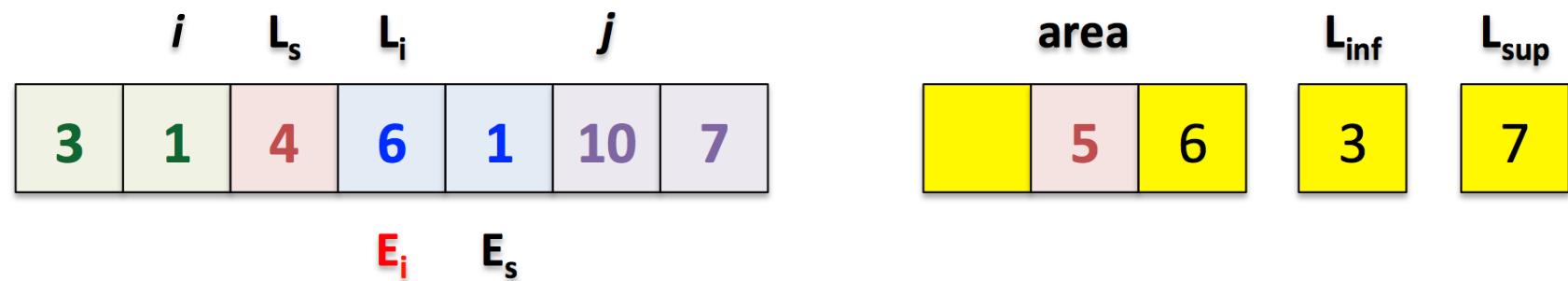
QuickSort Externo



QuickSort Externo



QuickSort Externo



QuickSort Externo

i	L_s	L_i	j
3	1	4	5

1 10 7

$$E_i = E_s$$

area	L_{inf}	L_{sup}
6	3	7

QuickSort Externo



QuickSort Externo

- Seja n o número de registros a serem ordenados e seja b o tamanho do bloco de leitura ou gravação do SO
- Melhor caso: da ordem de n/b operações
 - Ocorre quando o arquivo de entrada já está ordenado
- Pior caso: da ordem de n^2/m operações
 - Ocorre quando as partições geradas possuem tamanhos inadequados: maior tamanho possível e vazio
 - A medida que n cresce, a probabilidade de ocorrência do pior caso tende a zero
- Caso médio: da ordem de $(n/b) \log (n/m)$ operações

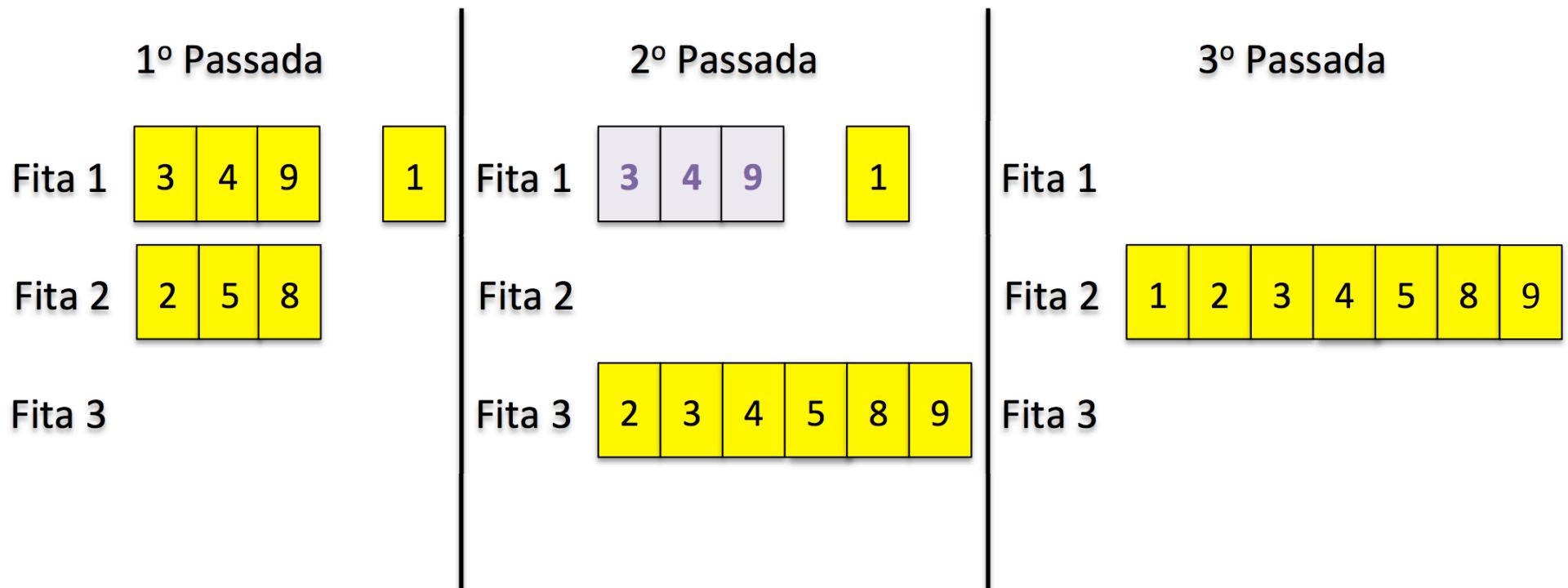
- Dada a sequência de números:

3 4 9 2 5 1 8

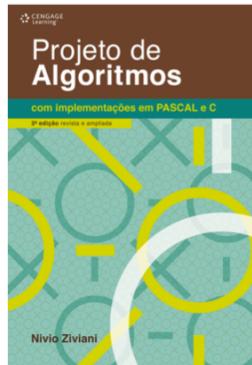
Ordene em ordem crescente utilizando o algoritmo de Intercalação Polifásica, apresentando a sequência dos números a cada passo. Considere um cenário em que a memória interna tem capacidade para armazenar dois números e com três unidades disponíveis de fita magnética.

Exercício

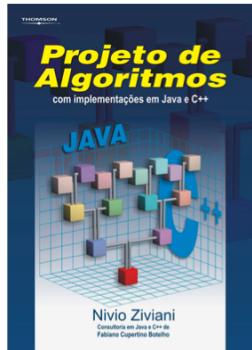
- Intercalação Polifásica:



Leitura Recomendada



ZIVIANI, N. Projeto de Algoritmos com Implementações em Pascal e C. 3^a Edição.
Cengage Learning, 2010. **Seção 4.2**



ZIVIANI, N. Projeto de Algoritmos com Implementações em Java e C++. Thomson Learning, 2007. **Seção 4.2**