Algoritmos e Estruturas de Dados III

Aula 10.2 - Casamento de Padrões - KMP

Prof. Hayala Curto 2022



Roteiro do Conteúdo



Casamento de Padrões

- Introdução
- Força Bruta
- KMP

KMP

- Algoritmo de Knuth-Morris-Pratt
- Inventado por Donald Knuth e Vaughan Pratt e independentemente por James Morris em 1977, embora os três tenham-no publicado conjuntamente.
- Introduz o conceito de autômato de estados e é usado para busca de cadeias.
- Procura a ocorrência de uma palavra W dentro de um "texto" S
- Quando aparece uma diferença, a palavra tem em si a informação necessária para determinar onde começar a próxima comparação

KMP

Características

Há fase de pré-processamento – complexidade O(m)

Comparações são feitas da esquerda para direita

Complexidade é O(m + n), m = tam. padrão, n = tam. texto

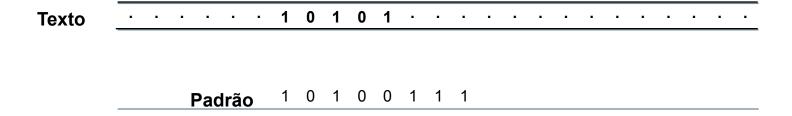
São realizadas 2*n* – 1 comparações no máximo

Análise

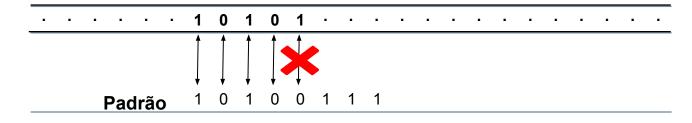
Melhor caso: O(m)

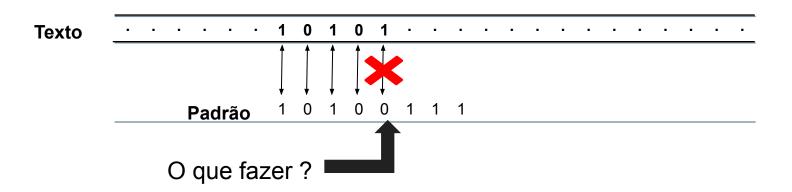
Pior caso: O(m + n)

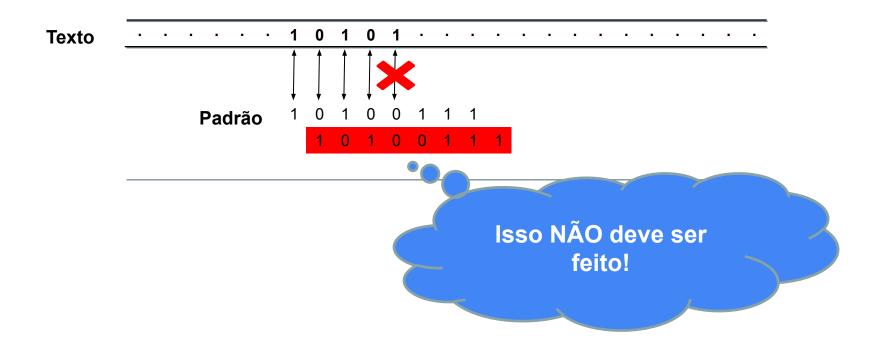
Caso médio: O(n)



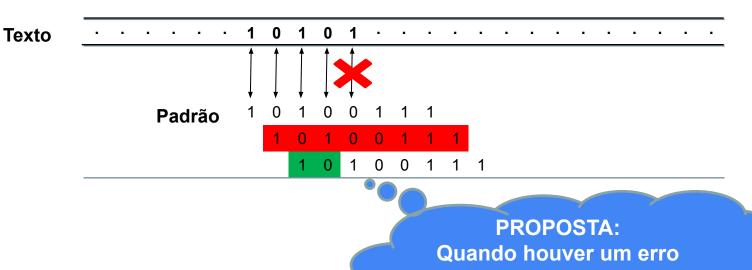








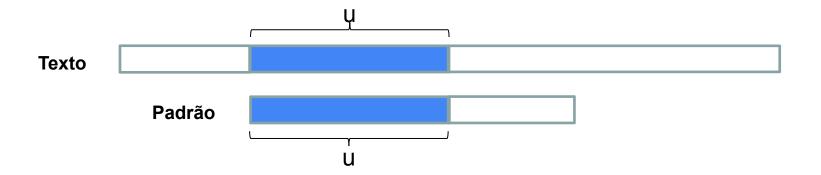


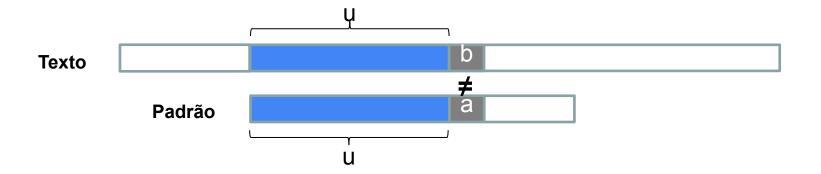


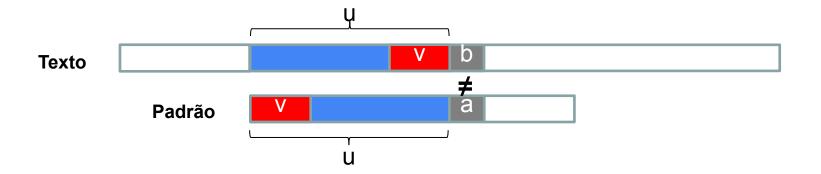
nunca retroceder no texto,

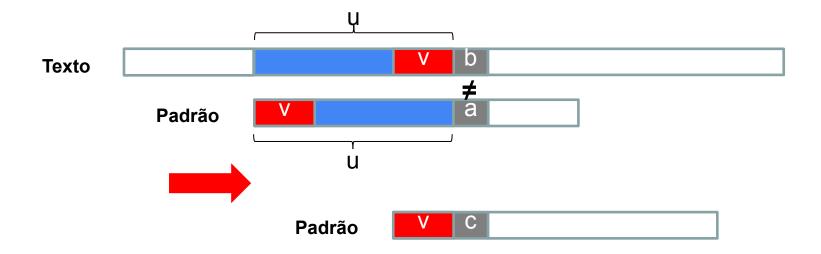
apenas no padrão!!!

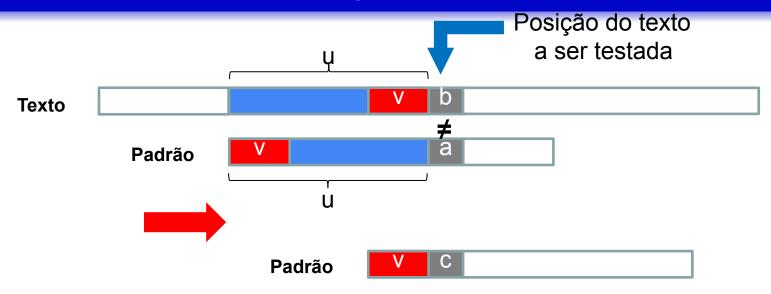
Texto			
	Padrão		

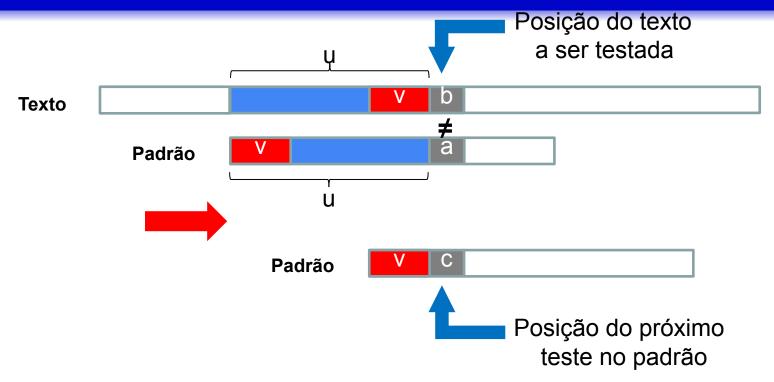


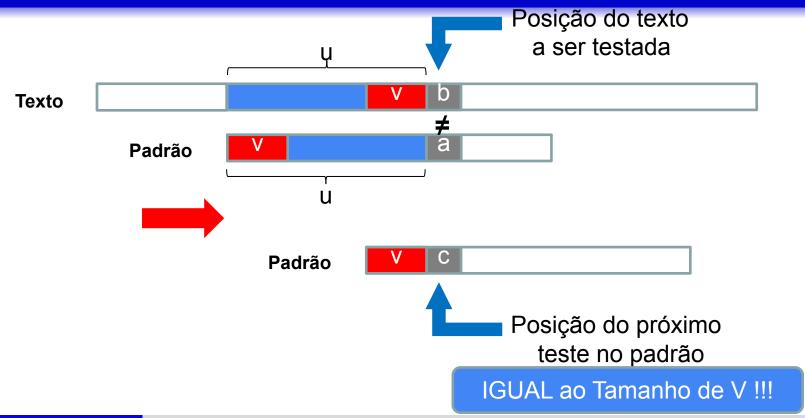




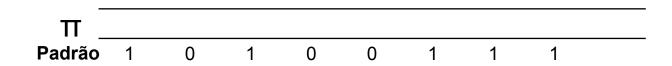


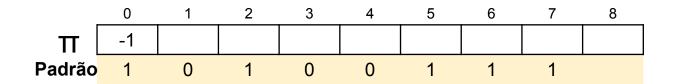


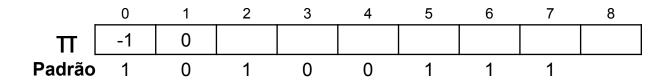


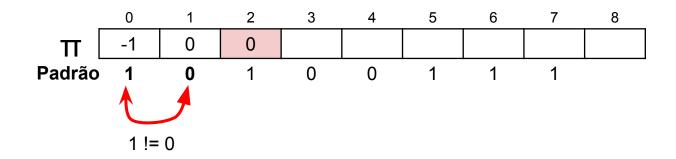


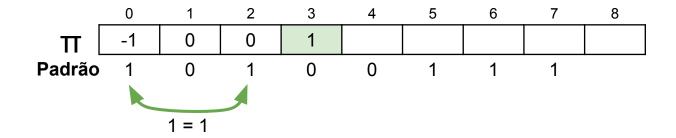
- Definição Função de Prefixo <u>Básica</u> (π)
 - π(j) = tamanho do maior prefixo v do padrão que é sufixo do texto antes do ponto de erro j

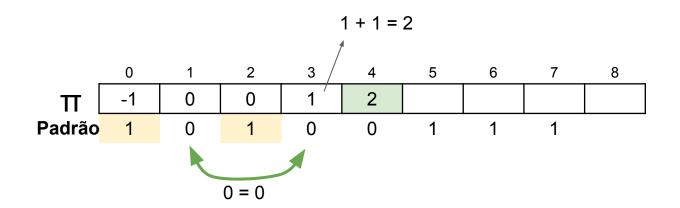


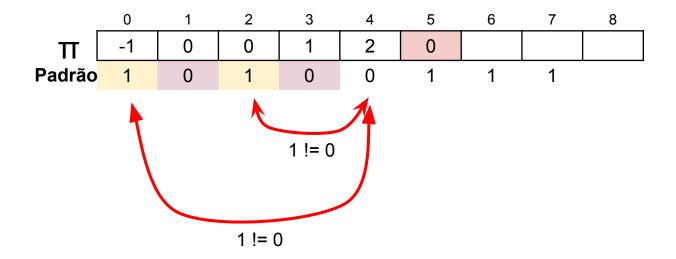


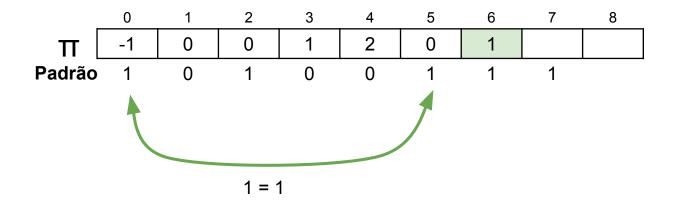


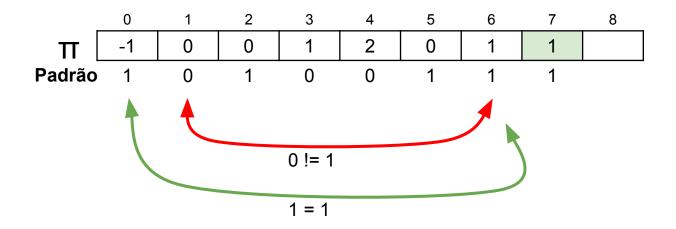


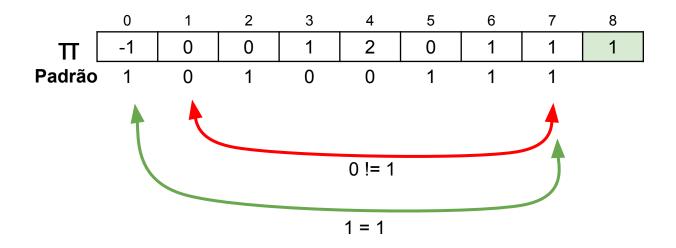


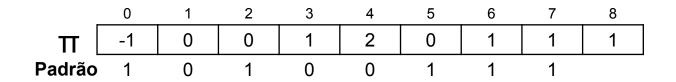




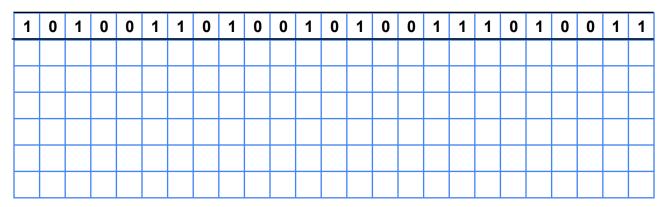


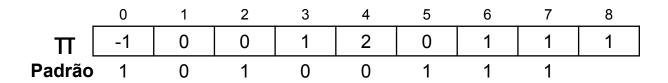


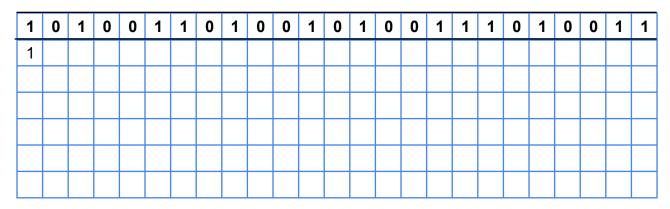


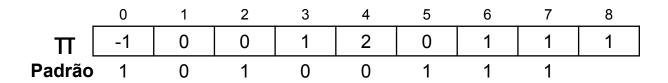


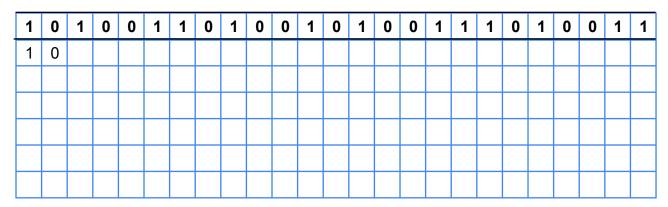
Texto

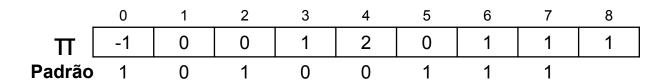


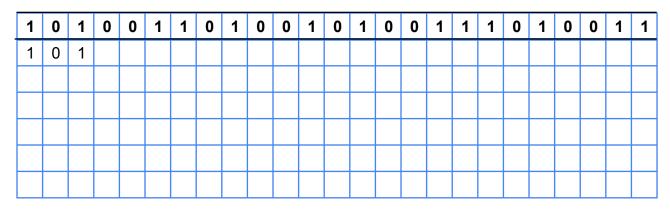


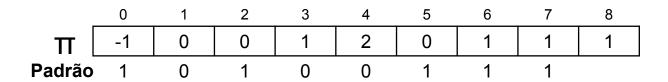


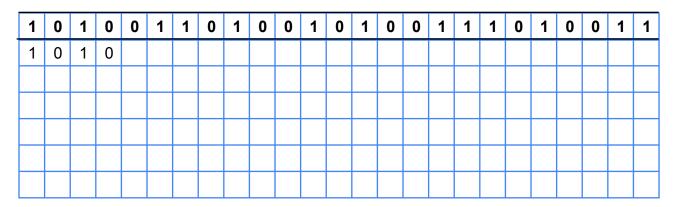


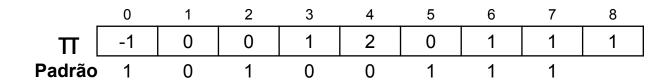


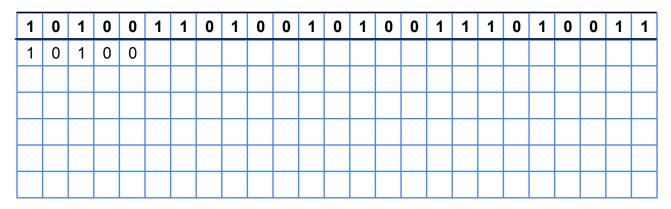


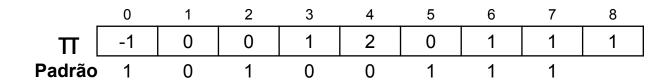


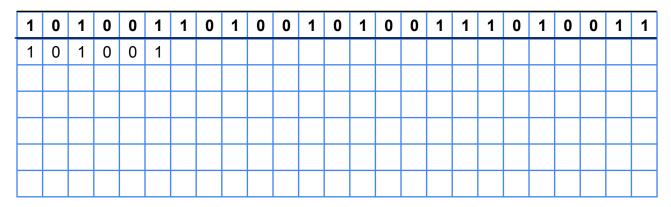


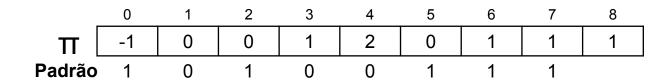


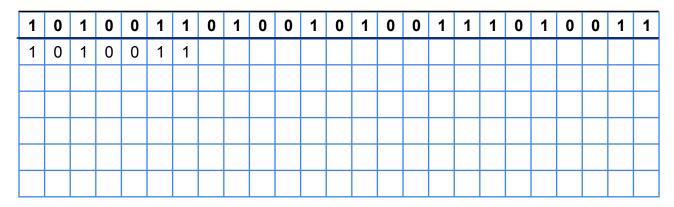


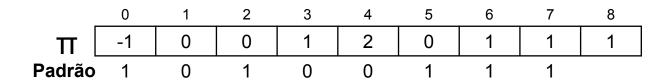


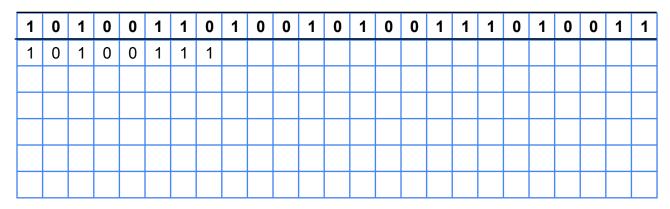


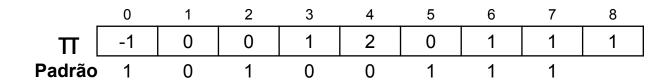


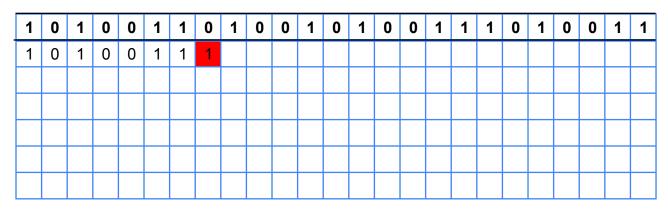


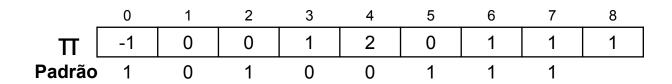


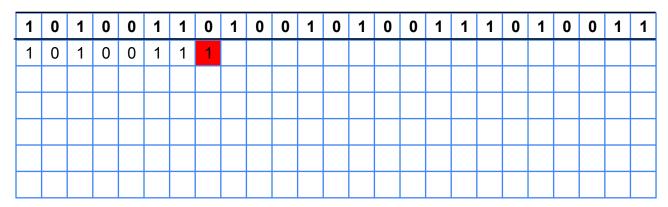


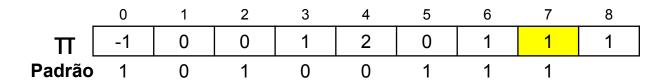


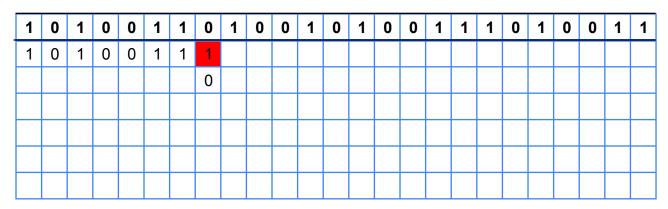


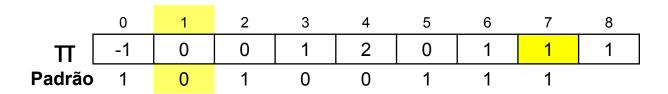


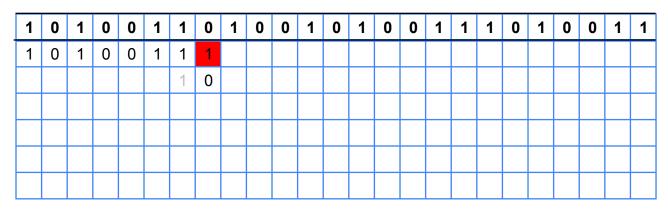


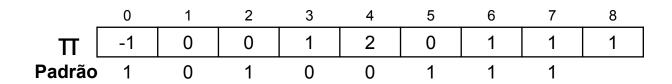


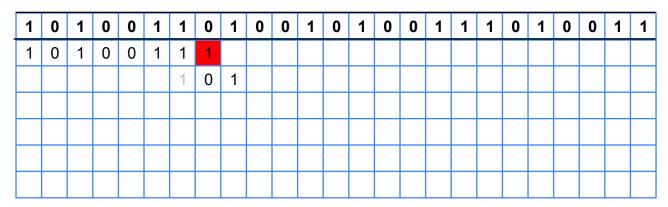


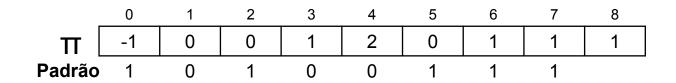


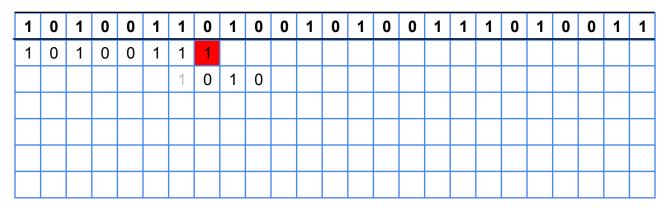


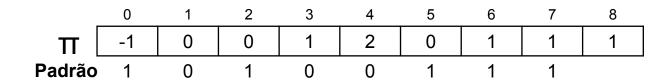


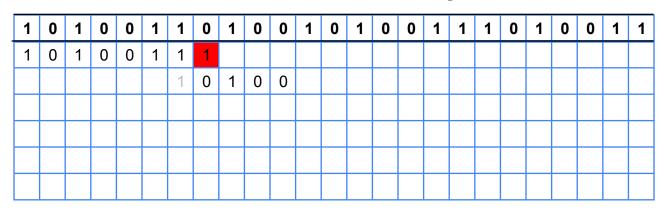


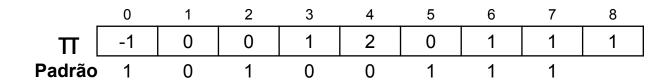


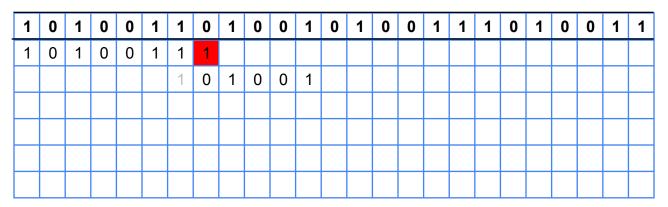


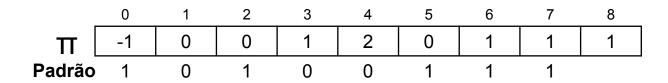


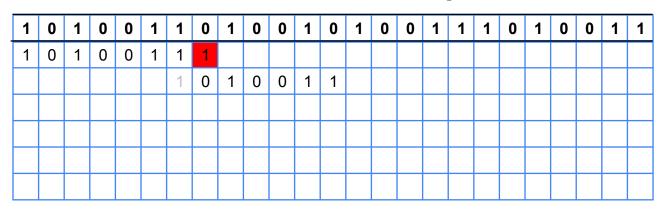


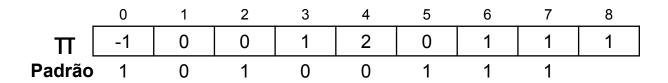


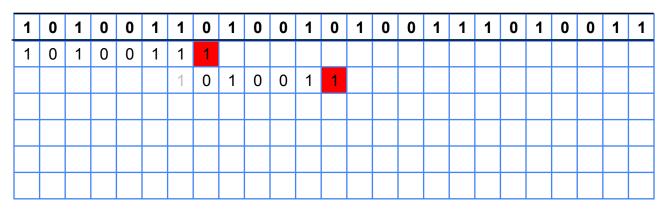


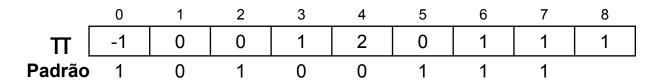


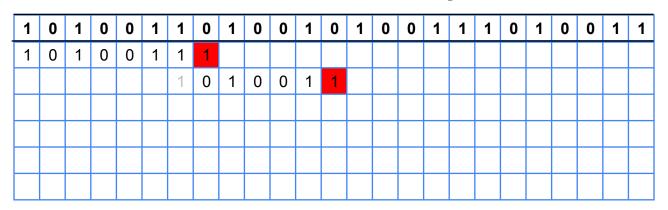


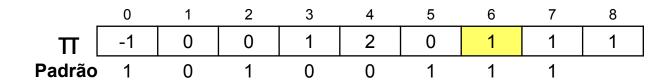


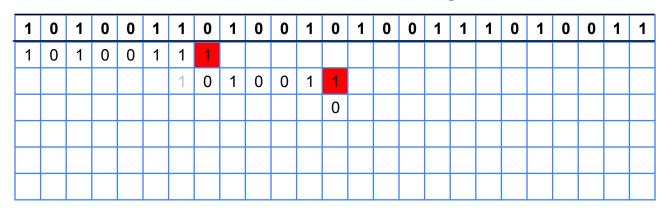


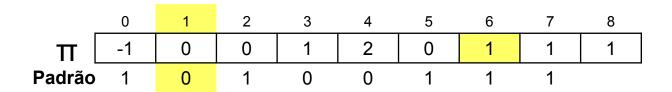


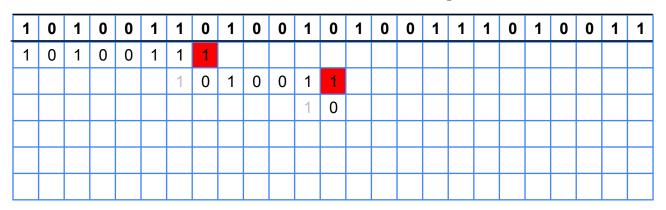


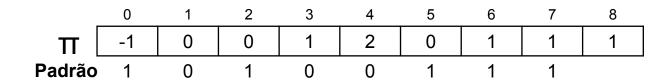


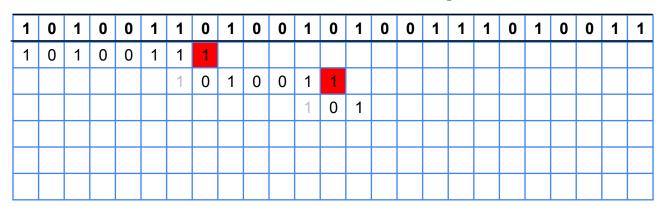


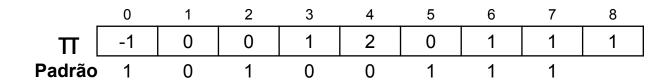


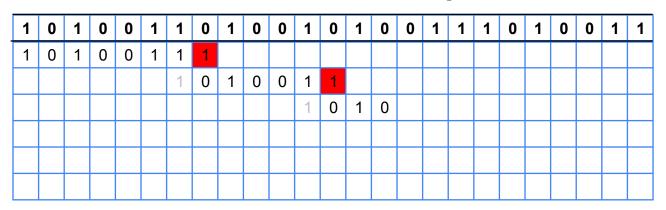


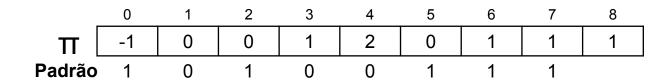


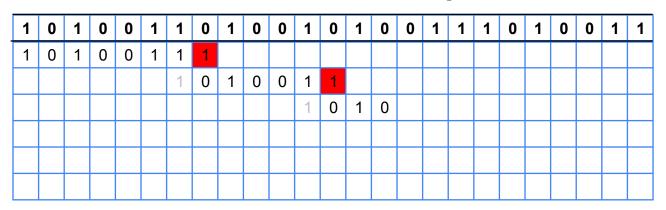


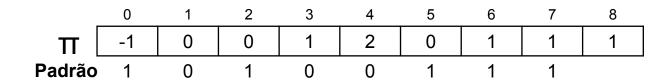


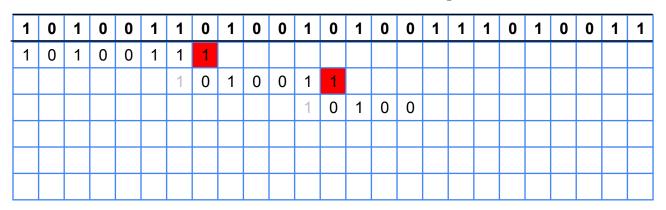


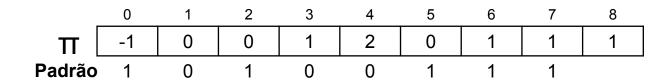


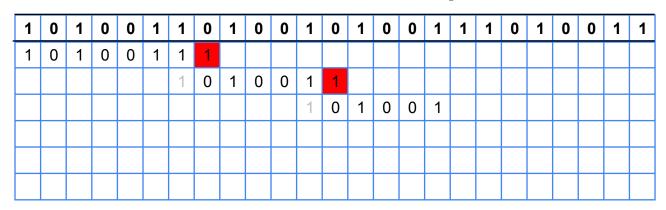


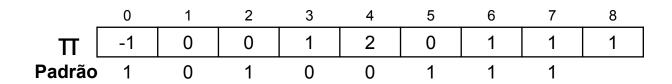


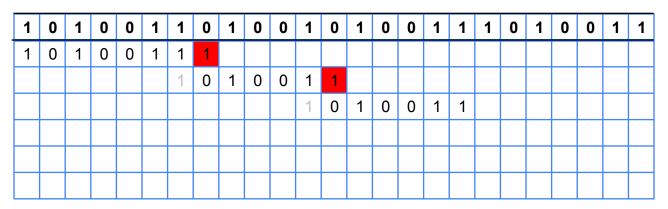


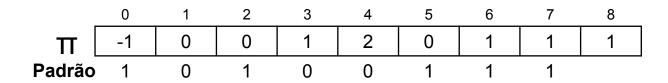


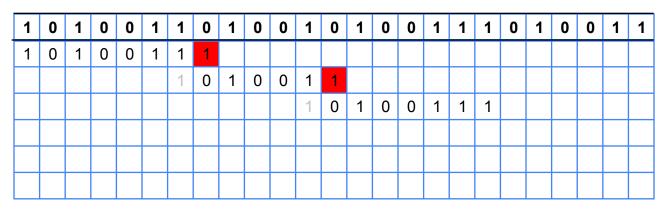


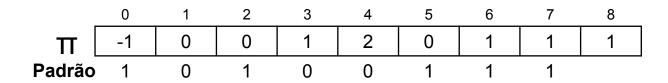


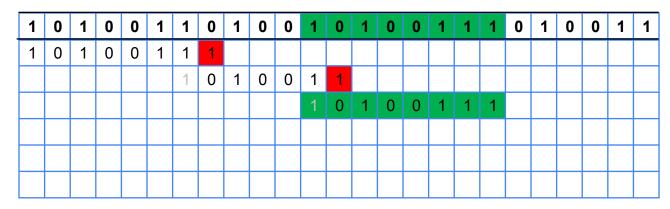










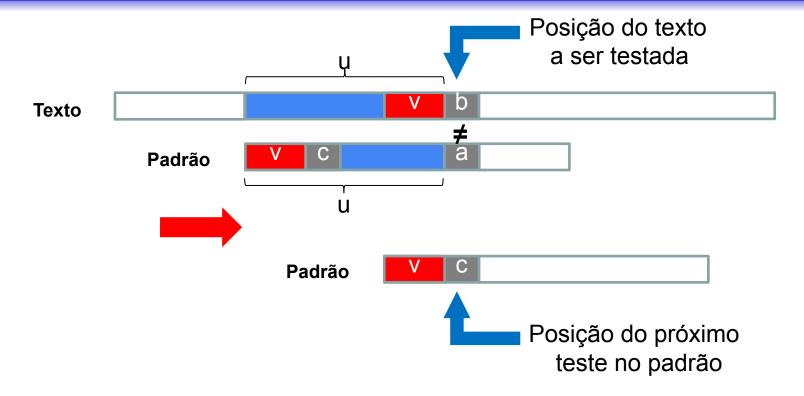


Comp.
8
6
7
21

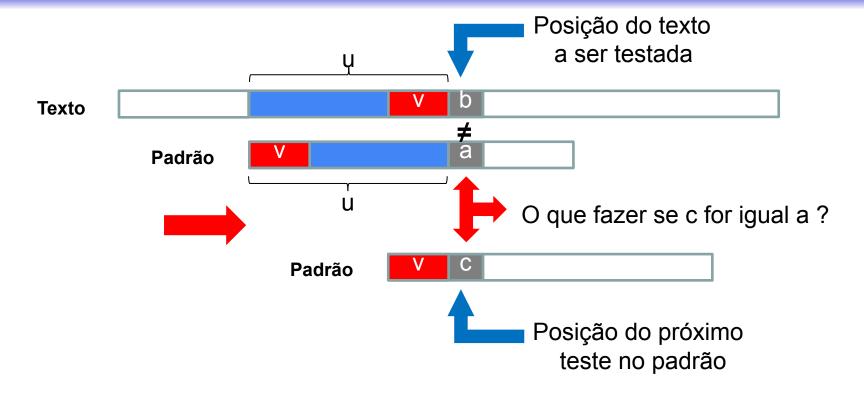
	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Padrão	1	0	1	0	0	1	1	1	-

https://cmps-people.ok.ubc.ca/ylucet/DS/KnuthMorrisPratt.html http://jovilab.sinaapp.com/visualization/algorithms/strings/kmp http://whocouldthat.be/visualizing-string-matching/

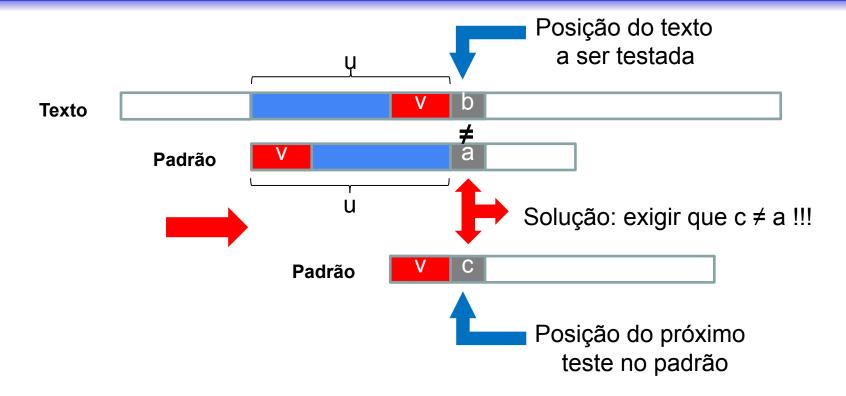
KMP – Função de Prefixo Básica



KMP – Função de Prefixo Básica

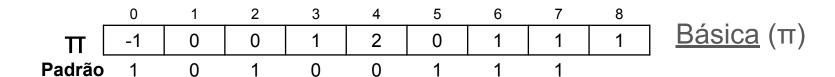


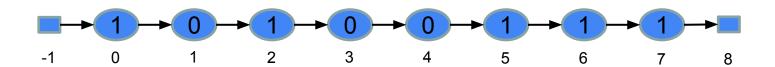
KMP – Função de Prefixo Melhorada

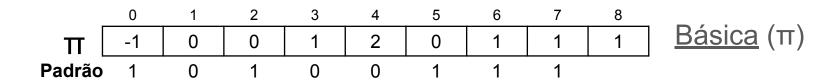


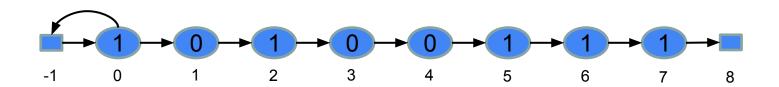
KMP – Função de Prefixo

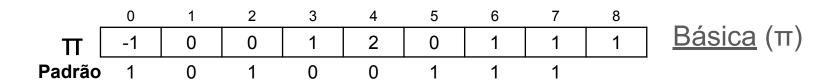
- Definição Função de Prefixo Melhorada (π)
 - π'(j) = tamanho do maior prefixo v do padrão que é sufixo do texto antes do ponto de erro j e é sucedido por símbolo diferente do que gerou o erro

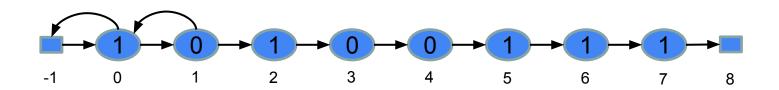


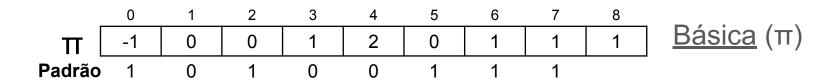


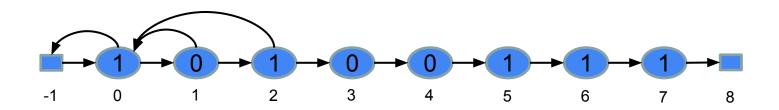


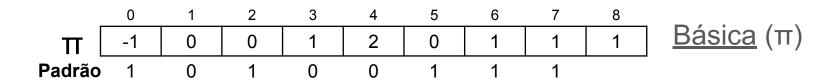


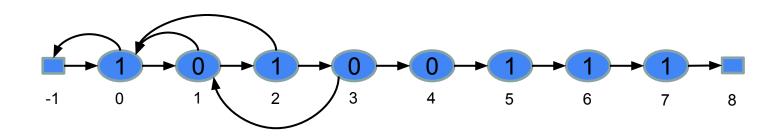


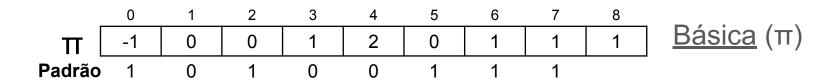


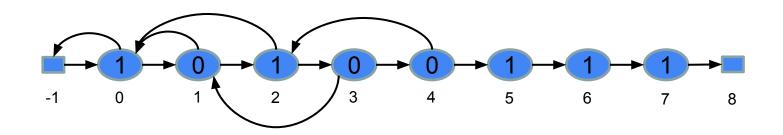


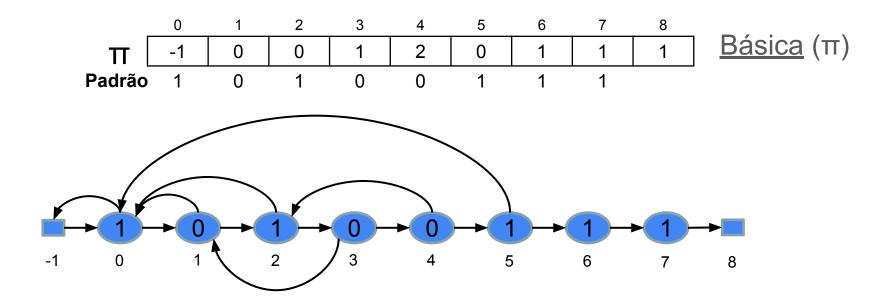


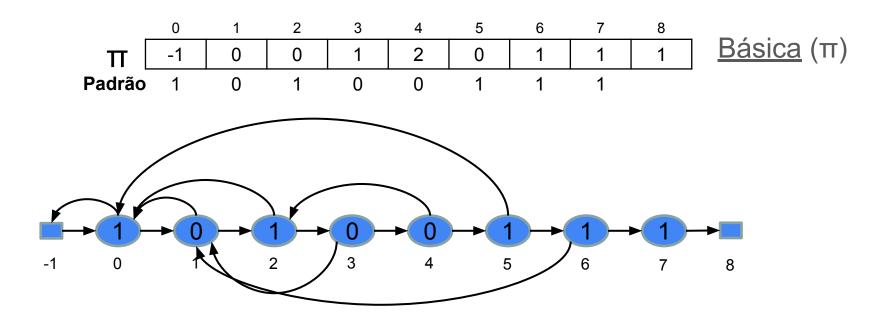


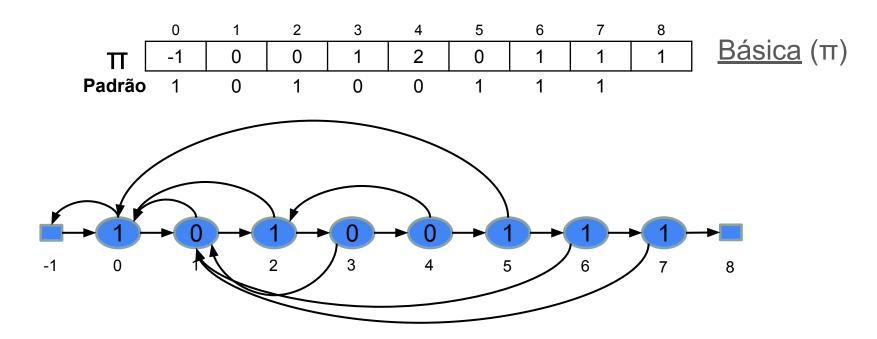


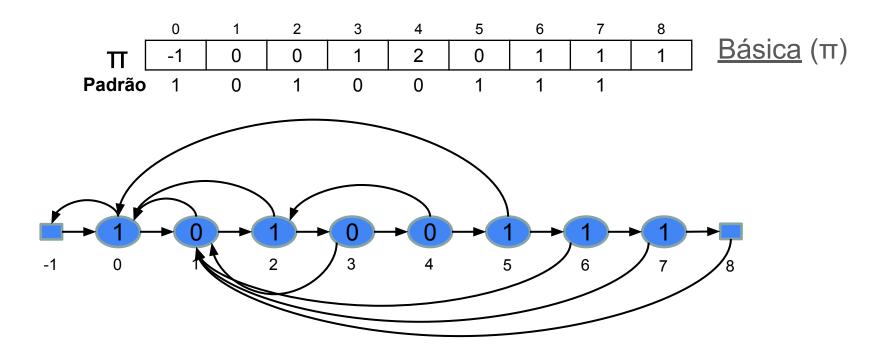


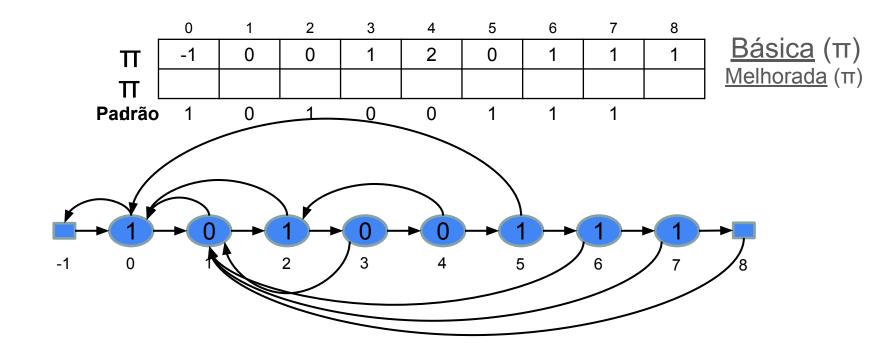


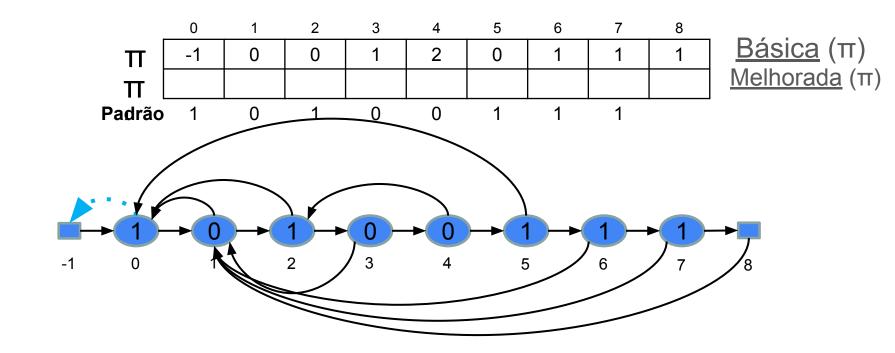


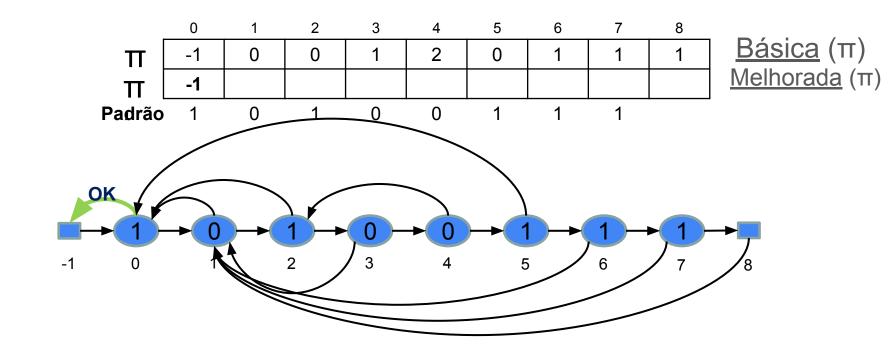


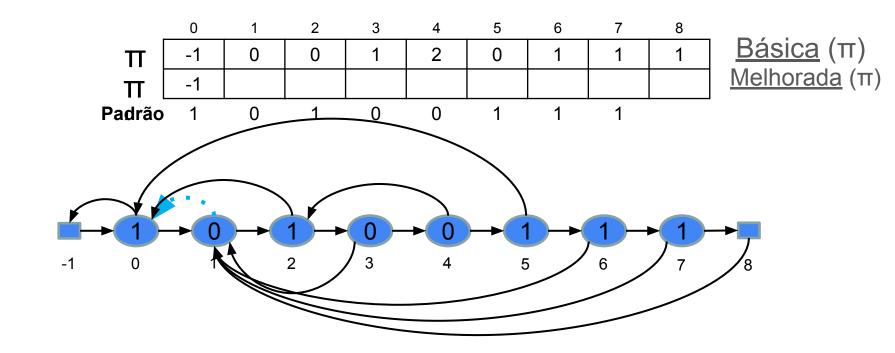


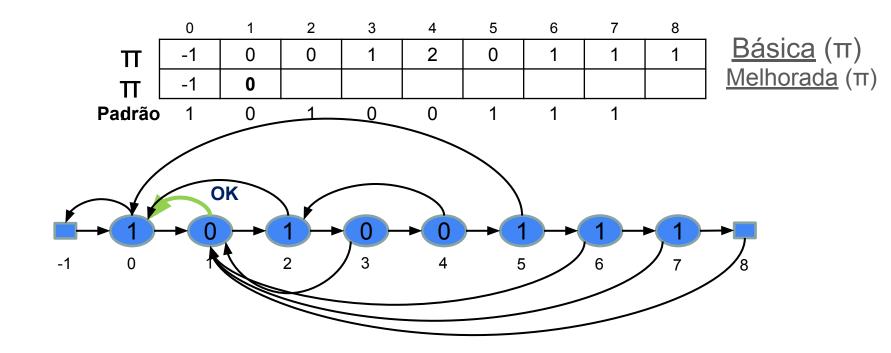


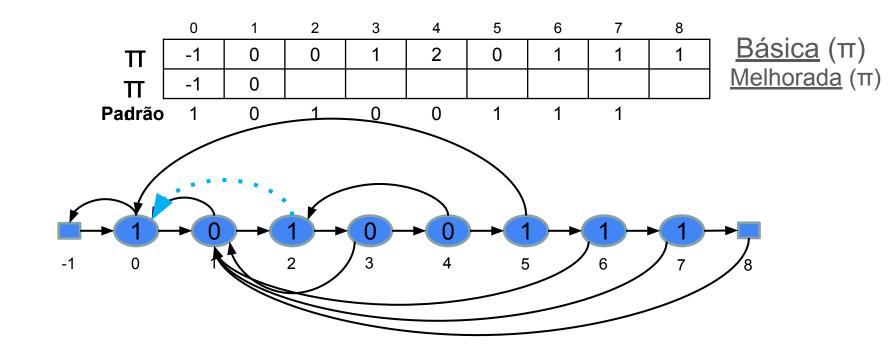


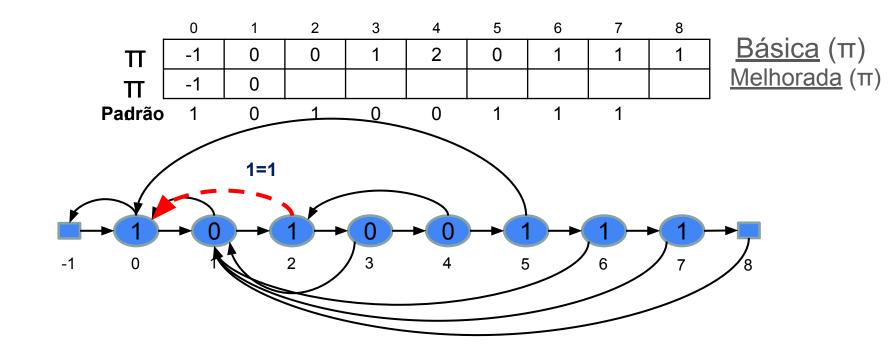


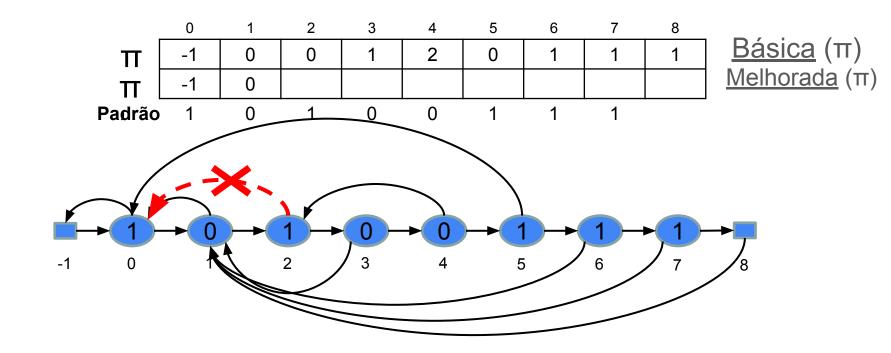


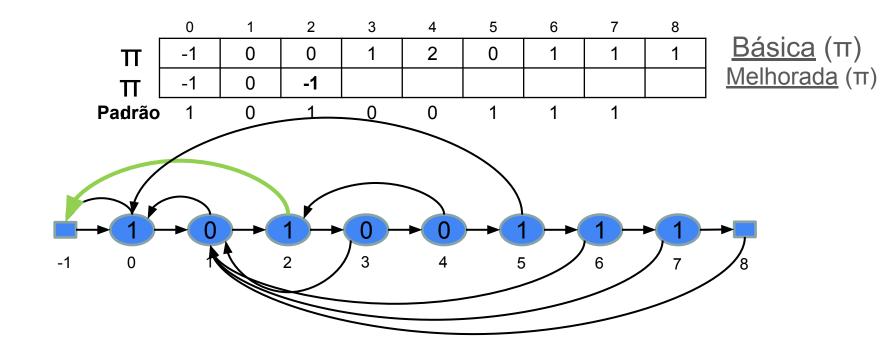


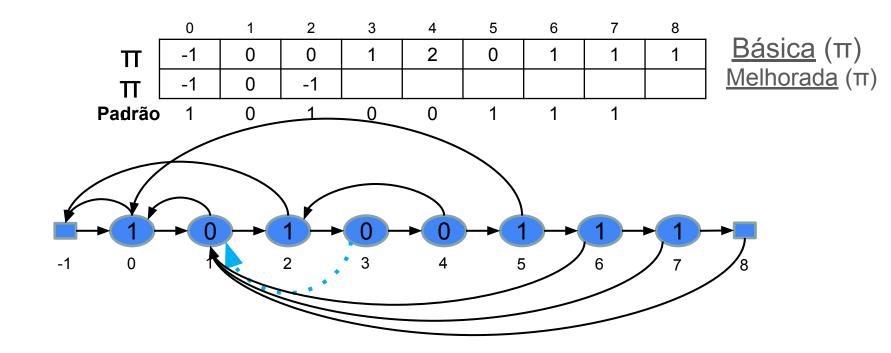


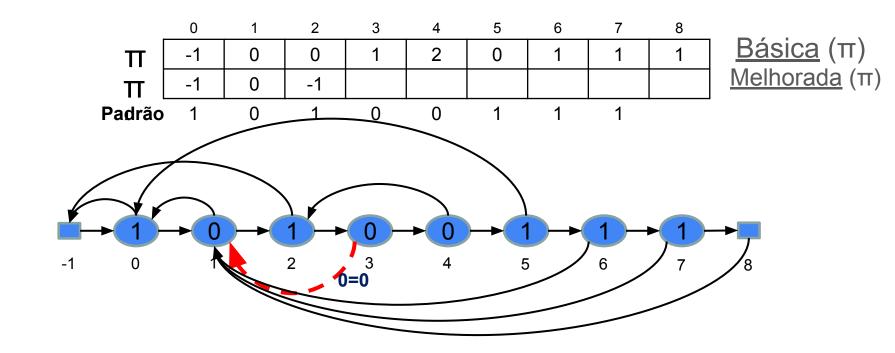


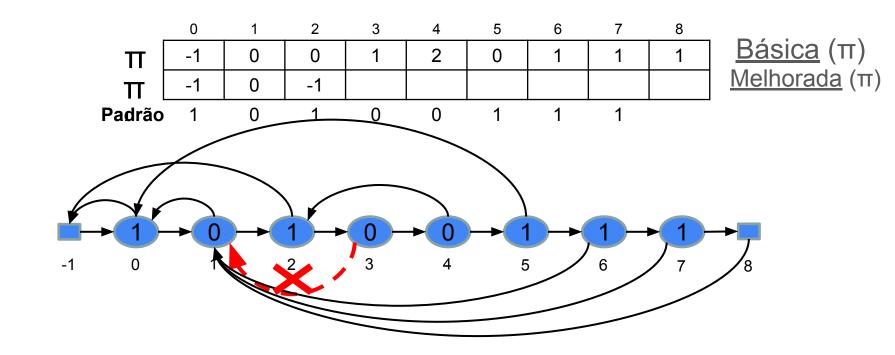


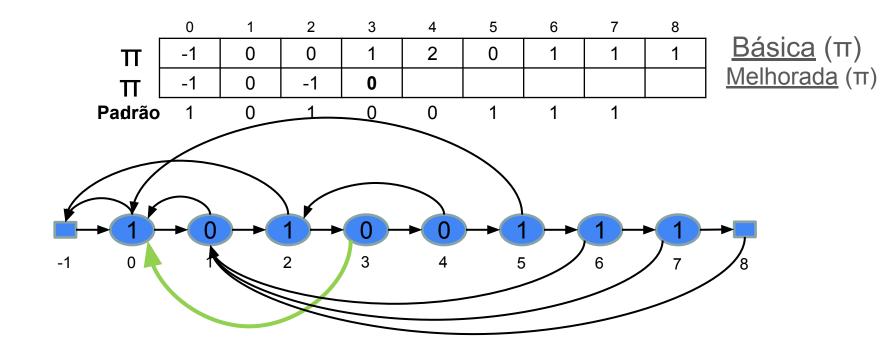


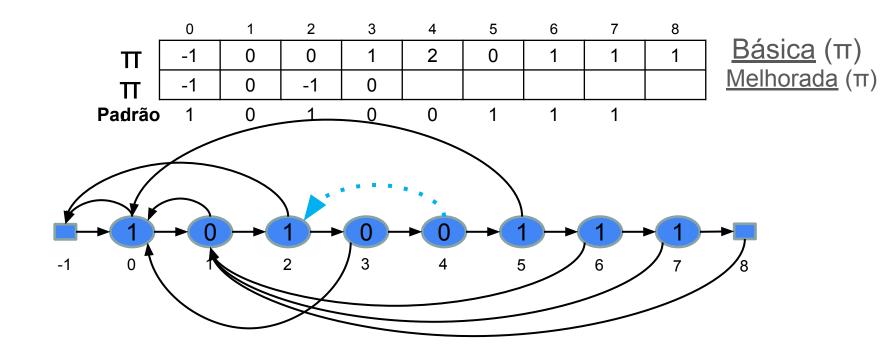


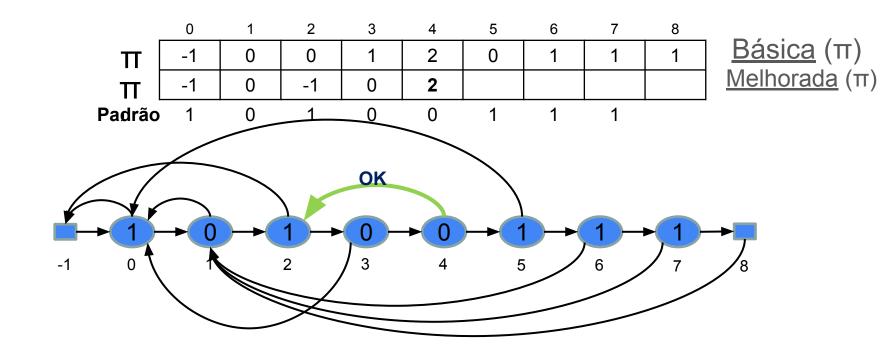


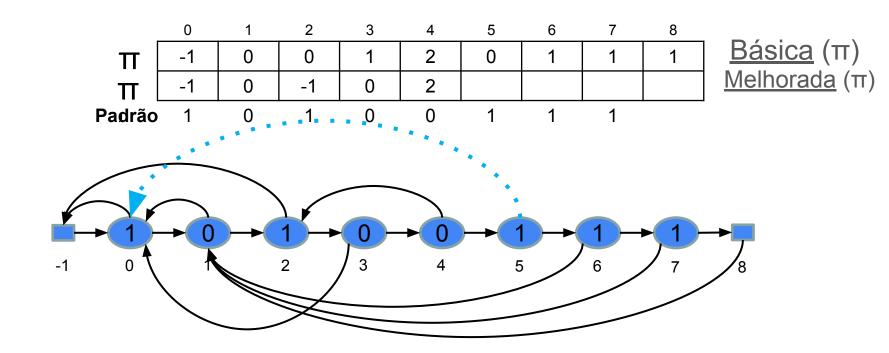


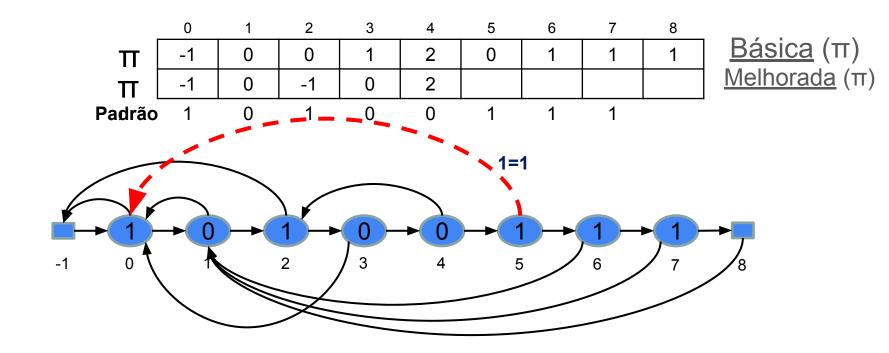


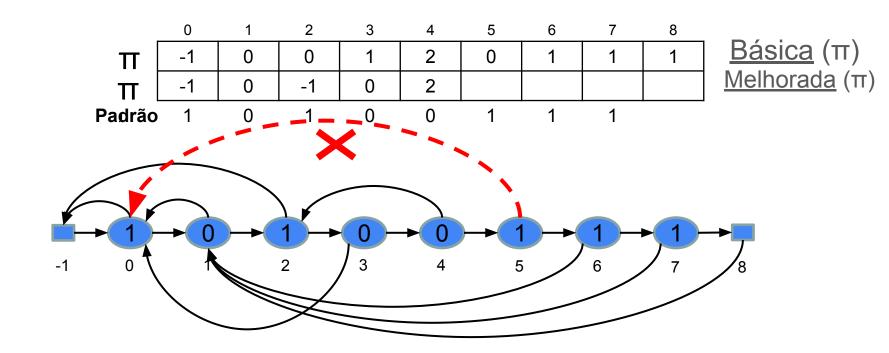


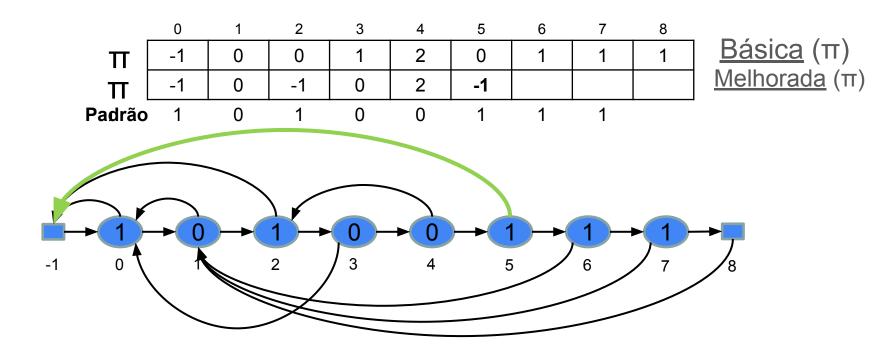


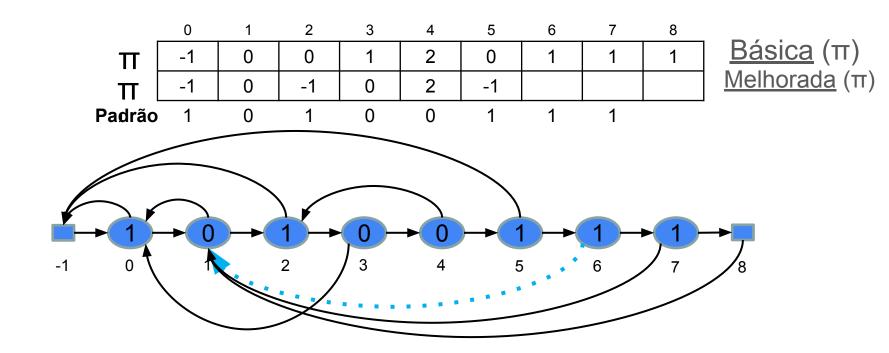


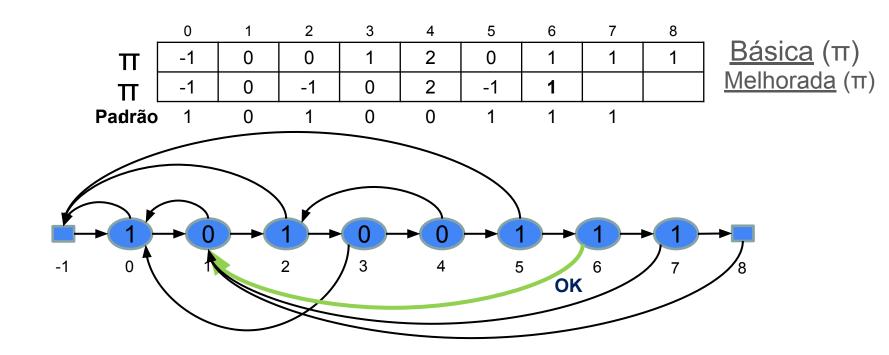


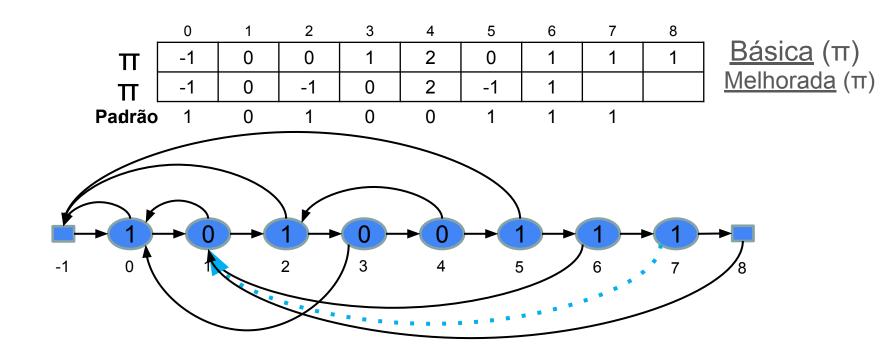


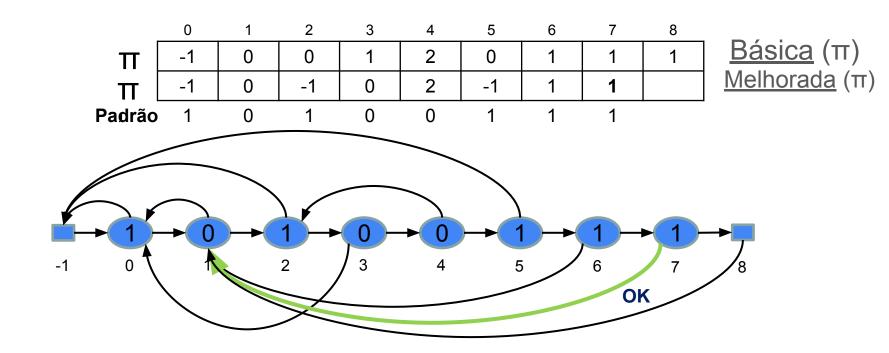


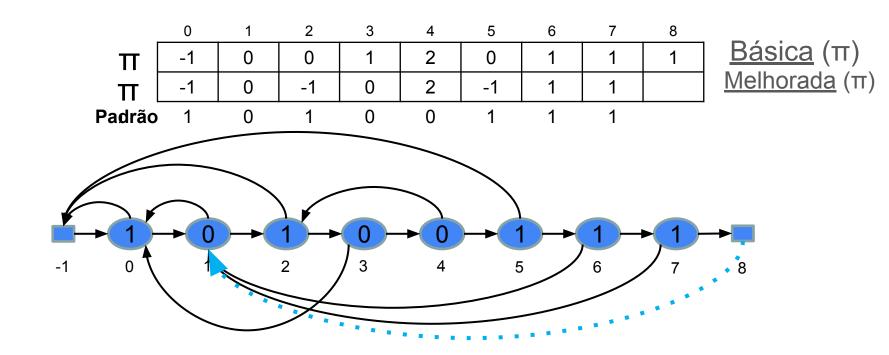


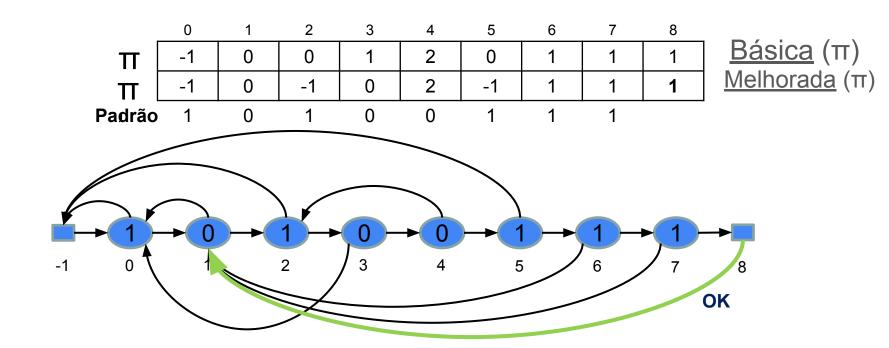


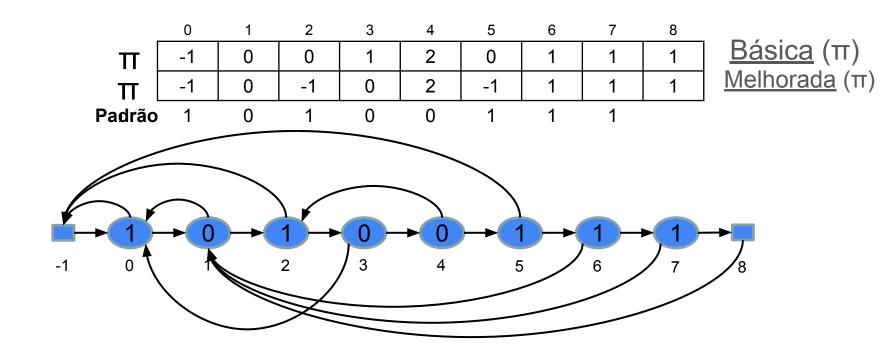






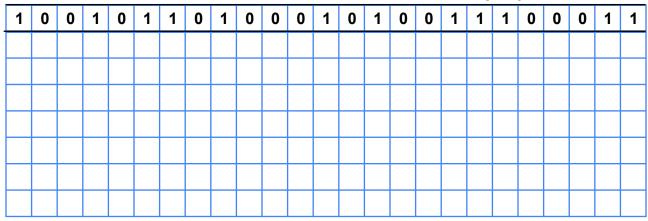






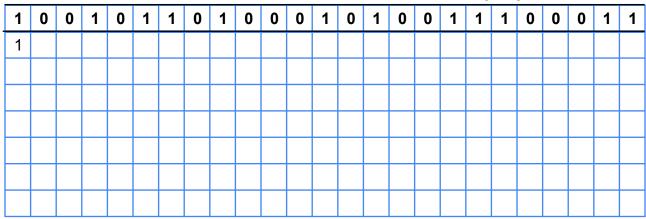
Usando Função de Prefixo <u>Básica</u> (π)

Texto



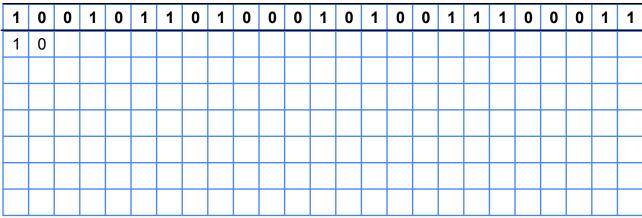
	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



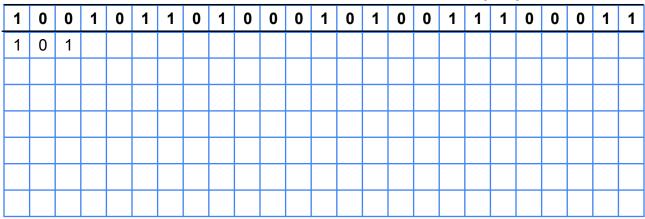
	0	1	2	3	4	5	6	7	8
Π	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



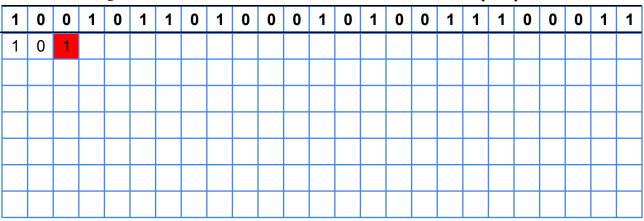
	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



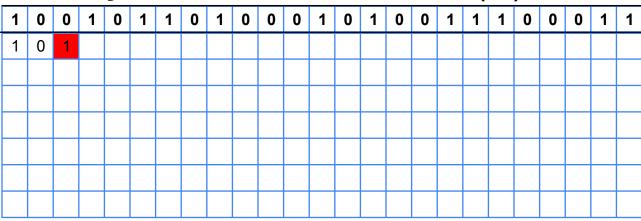
	0	1	2	3	4	5	6	7	8
Π.	-1	0	0	1	2	0	1	1	1
π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



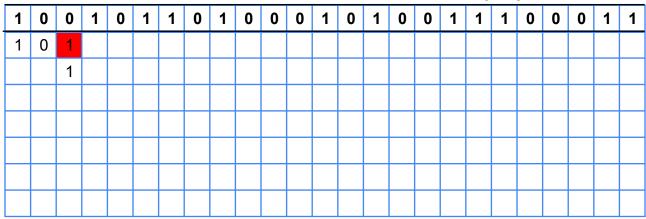
	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



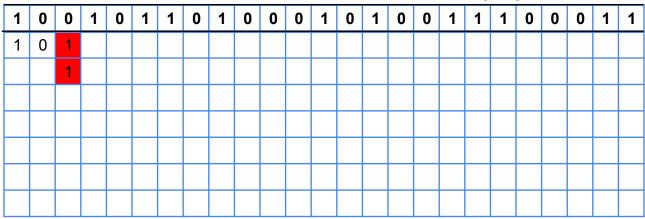
	0	1	2	3	4	5	6	7	8
Π.	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



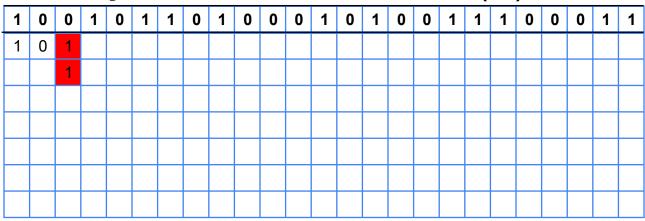
	0	1	2	3	4	5	6	7	8
Π.	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



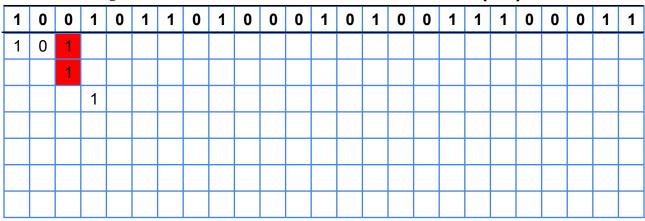
	0	1	2	3	4	5	6	7	8
Π	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



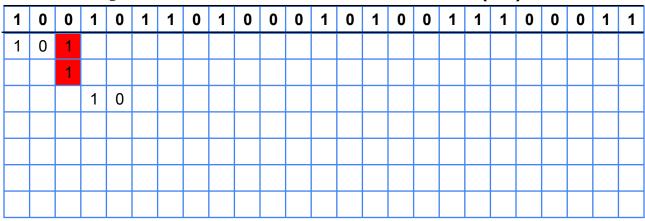
_	0	1	2	3	4	5	6	7	8
Π	-1	0	0	1	2	0	1	1	1
Π ' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



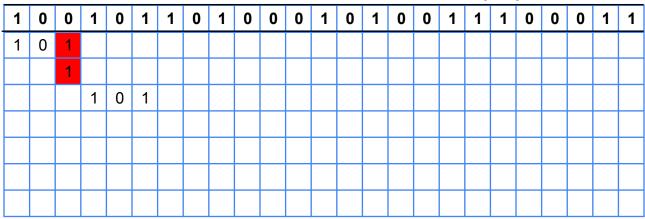
	0	1	2	3	4	5	6	7	8
Π.	-1	0	0	1	2	0	1	1	1
π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



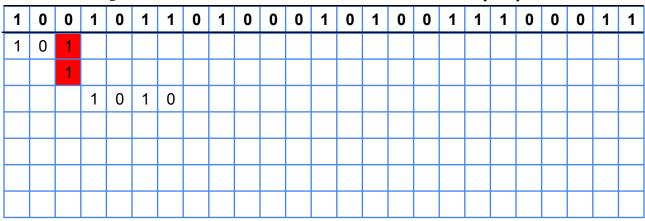
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
- Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



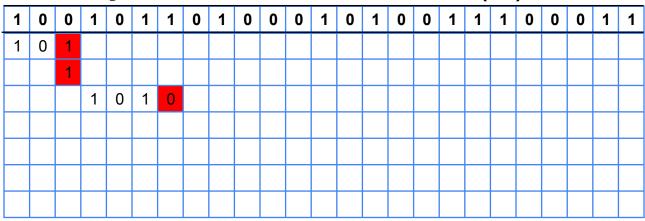
	0	1	2	3	4	5	6	7	8
Π.	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



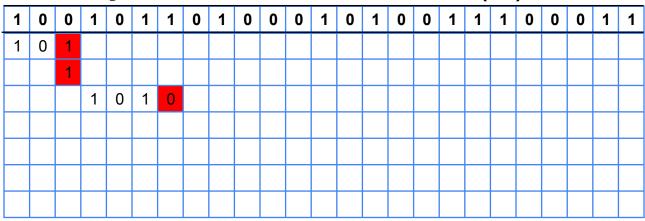
	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



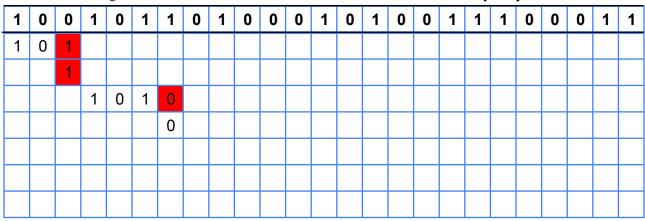
	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



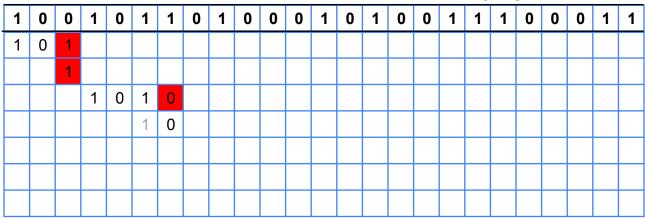
_	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
Π ' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



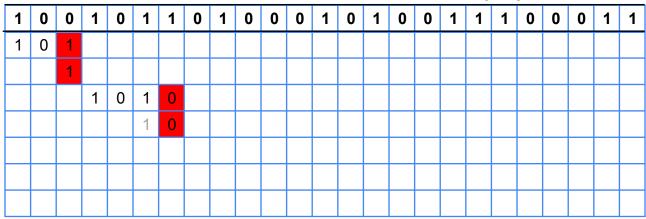
	0	1	2	3	4	5	6	7	8
Π	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



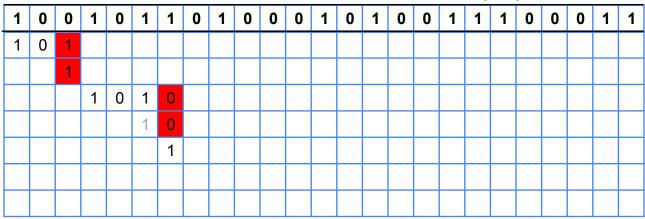
	0	1	2	3	4	5	6	7	8
Π.	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



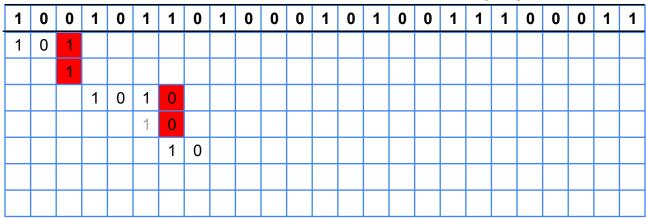
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



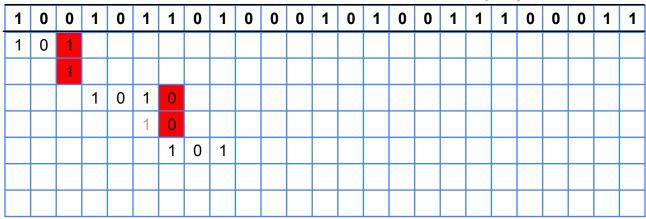
	0	1	2	3	4	5	6	7	8
Π.	-1	0	0	1	2	0	1	1	1
Π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



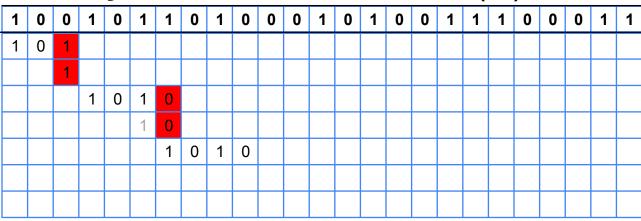
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



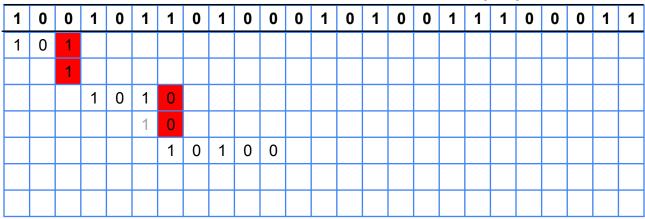
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



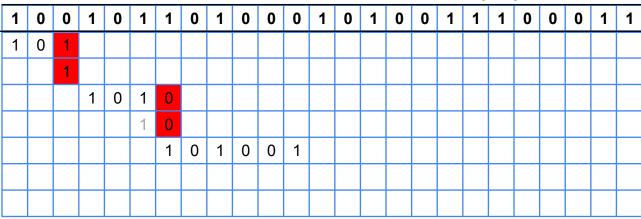
_	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



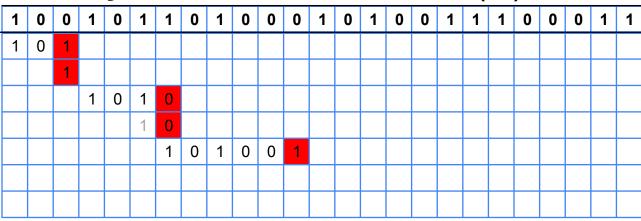
_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



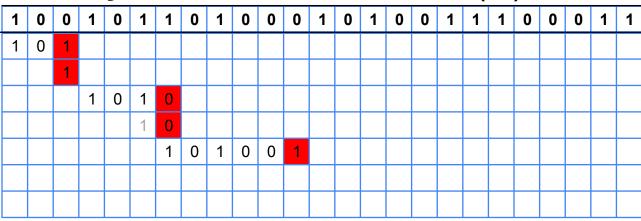
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



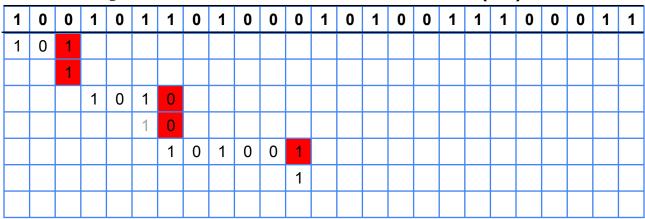
_	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



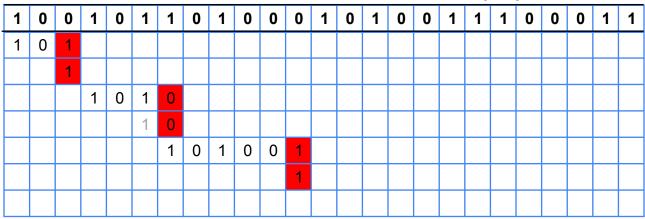
_	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



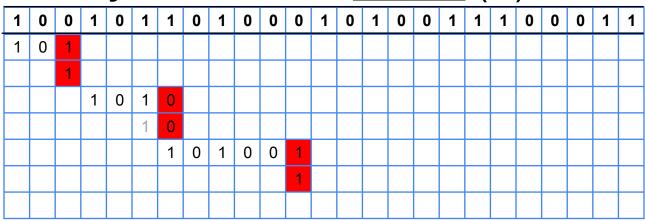
	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



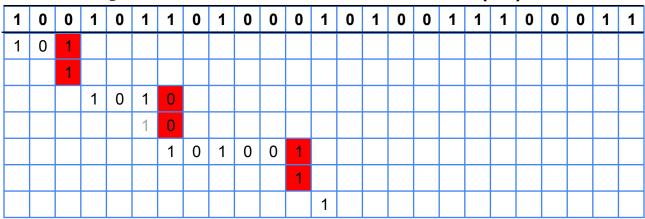
_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



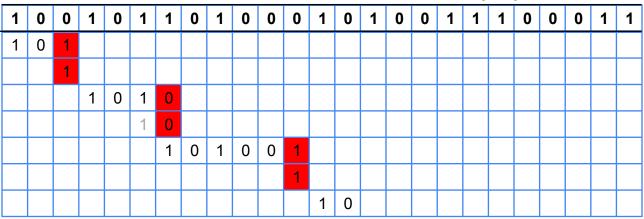
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Π ' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



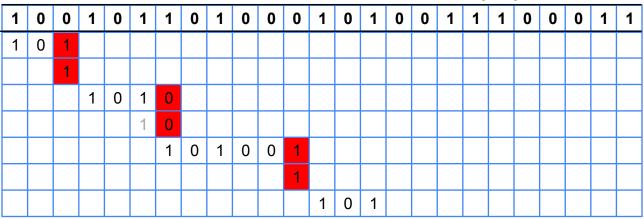
	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



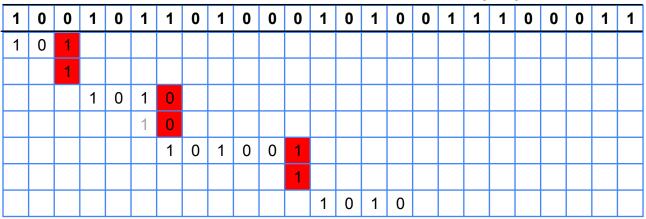
_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



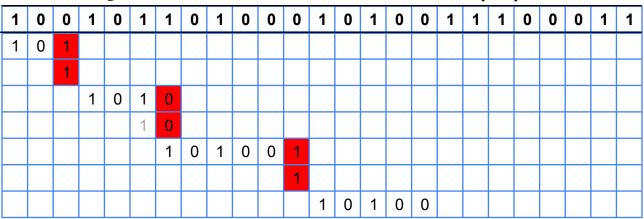
_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



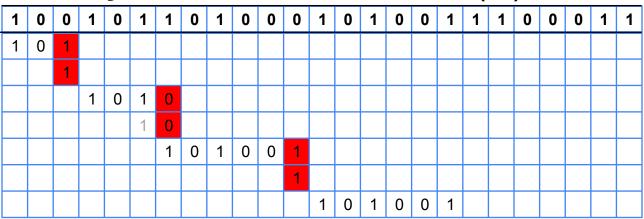
_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



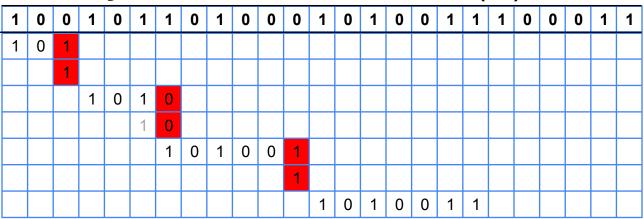
_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



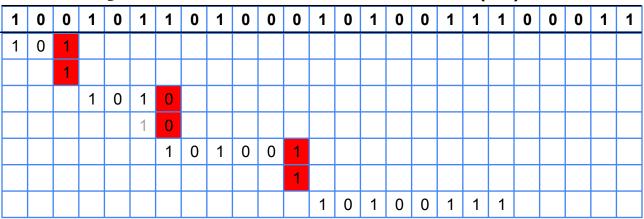
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



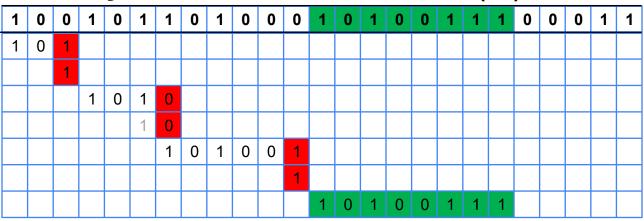
_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)



_	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

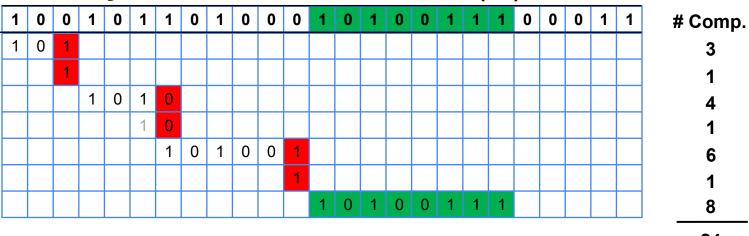
Usando Função de Prefixo <u>Básica</u> (π)



_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo <u>Básica</u> (π)

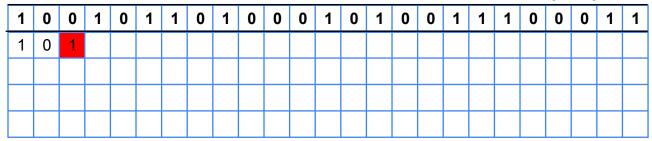
Texto Padrão



24

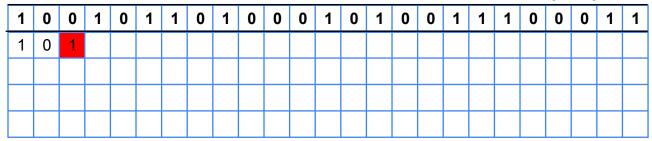
	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π'	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')



_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Π ' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')



_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Π	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')

1	0	0	1	0	1	1	0	1	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	1
1	0	1																						
			1	0	1	0																		

_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Π ' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')

1	0	0	1	0	1	1	0	1	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	1
1	0	1																						
			1	0	1	0																		

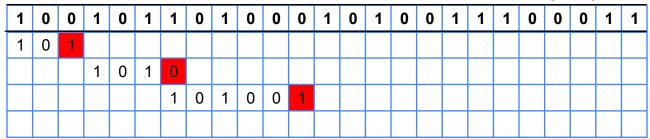
_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Π	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')

1	0	0	1	0	1	1	0	1	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	1
1	0	1																						
			1	0	1	0																		
						1	0	1	0	0	1													

_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Π	-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')



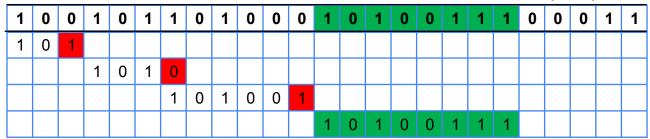
_	0	1	2	3	4	5	6	7	8
Π_{\cdot}	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')

1	0	0	1	0	1	1	0	1	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	1
1	0	1																						
			1	0	1	0																		
						1	0	1	0	0	1													
												1	0	1	0	0	1	1	1					

_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

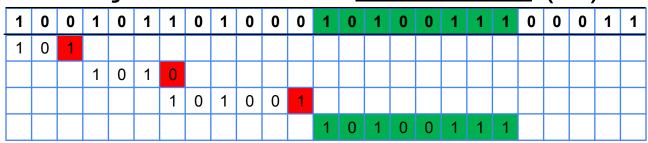
Usando Função de Prefixo Melhorada (π')



_	0	1	2	3	4	5	6	7	8
π	-1	0	0	1	2	0	1	1	1
Π ' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Usando Função de Prefixo Melhorada (π')

Texto Padrão



6 8

Comp.

21

_	0	1	2	3	4	5	6	7	8
π. [-1	0	0	1	2	0	1	1	1
π' [-1	0	-1	0	2	-1	1	1	1
Padrão	1	0	1	0	0	1	1	1	

Algoritmo

```
CalculaFunçãoPrefixo(P)
M \leftarrow \text{Tamanho}(P);
\pi[0] = -1;
\mathbf{for} \ (i = 0, j = -1; i <= M; i ++, j ++, \pi[\ i\ ] = j\ )
\mathbf{while} \ \ (j >= 0 \&\& P[\ i\ ] != P[\ j\ ])
j = \pi[\ j\ ];
\mathbf{return} \ \pi;
```

Algoritmo

CalculaFunçãoPrefixoMelhorada(P)

```
M \leftarrow \text{Tamanho}(P);

\pi'[0] = -1;

\text{for } (i = 0, j = -1; i \le M; i + +, j + +,

\pi'[i] = (P[i] == P[j]) ? \pi'[j] : j)

\text{while } (j >= 0 && P[i] != P[j])

j = \pi'[j];

\text{return } \pi';
```

Algoritmo

```
KMP(T, P)
       N \leftarrow \text{Tamanho}(T);
       M \leftarrow \text{Tamanho}(P);
       \pi' \leftarrow \text{CalculaFuncaoPrefixoMelhorada(P)};
       for (i = 0, j = 0; i < N \&\& j < M; i++, j++)
              while (j \ge 0 \&\& T[i] != P[j])
                      j = \pi'[j];
       if (j == M) return i - M;
       else return i;
```