




# Tutoria - 06

 Lucas de Araújo | 18.2.4049

Questões

- Questão 1)
  - Heurística Ocorrência
  - Heurística Casamento
- Questão 2)
- Questão 3)
- Questão 4)
- Questão 5)
  - A)
  - B)

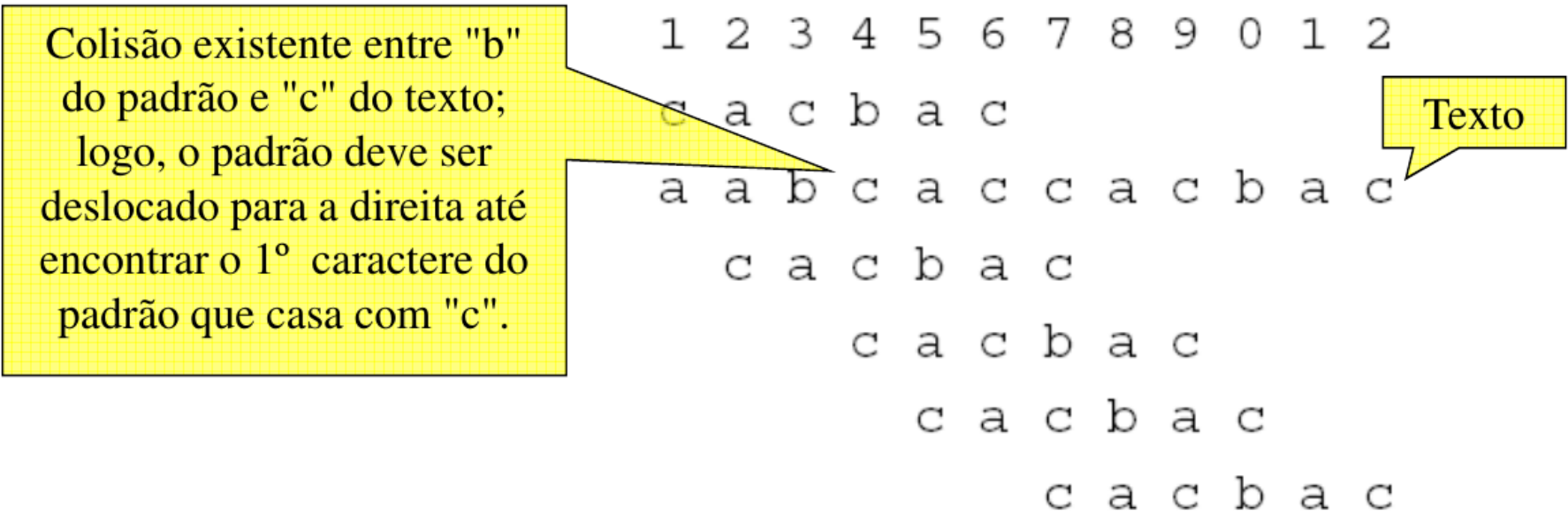
## Questões

### Questão 1)

#### Heurística Ocorrência

A heurística ocorrência alinha o caractere no texto que causou a colisão com o primeiro caractere no padrão, a esquerda do ponto de colisão, que casa com ele.

$$P = \{cacbac\}$$
$$T = \{aabcaccacbac\}$$



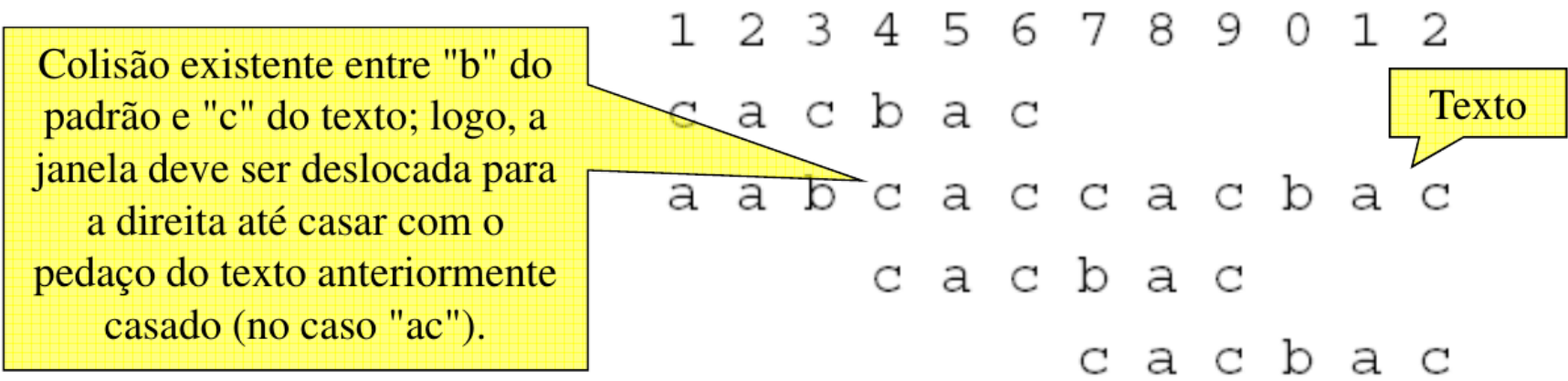
Caso o caractere não seja encontrado no restante do padrão após colisão, então o tamanho caminhado pelo deslocamento da janela será igual o tamanho do padrão.

**Observação:** Quanto maior o tamanho do alfabeto, melhor a heurística ocorrência se torna quando comparada a força bruta em termos de eficiência

#### Heurística Casamento

A heurística casamento faz com que, ao mover o padrão para direita, a janela em questão casa com o pedaço do texto anteriormente casado

$$P = \{cacbac\}$$
$$T = \{aabcaccacbac\}$$



O algoritmo Boyer-Moore decide qual das duas heurísticas deve seguir, escolhendo aquela que provoca o maior deslocamento do padrão. Esta escolha implica em realizar comparações para cada colisão que ocorrer, penalizando o desempenho do algoritmo com relação a tempo de processamento

Ao longo dos anos, várias propostas de simplificação surgiram, sendo que os melhores resultados foram obtidos por aquelas que consideraram apenas a heurística ocorrência

### Questão 2)

A tabela de deslocamento é definida da seguinte forma:

- O valor inicial do deslocamento para todo os caracteres do texto é igual a  $m$
- Em seguida, para os  $m - 1$  primeiros caracteres do padrão  $P$ , os valores do deslocamento são calculados pela seguinte regra:

$$d[x] = \min\{j \mid (j = m) \mid (1 \leq j < m \ \& \ P[m - j] = x)\}$$

- **Exemplo:** Para o padrão  $P = \{teste\}$ , os valores da tabela são:
  - $d["t"] = 1,$
  - $d["e"] = 3,$
  - $d["s"] = 2;$
  - $d[x] = 5$  (valor de  $m$ ) para todo caractere  $x$  do texto que não faça parte do padrão
- **Padrão  $P$ :** AMAFA
  - $d[a] = 2$
  - $d[m] = 3$
  - $d[f] = 1$

### Questão 3)

Fizemos a desconsideração do último caractere pois graças a isto, podemos deslocar a tabela ao sofrer colisão e alinhar o último caractere do texto com o último caractere do padrão. Dessa forma, de maneira consecutiva, podemos buscar o padrão no texto através desses alinhamentos

### Questão 4)

A simplificação apresentada por Sunday foi considerar a o último caractere na criação da tabela de deslocamentos durante a fase de pré-processamento e com isso, deslocar a tabela de deslocamentos relativo ao caractere do texto correspondente ao caractere após o último caractere do padrão (fim do padrão + 1)

Com esta simplificação, foi observável as seguintes melhorias:

- Redução de deslocamentos em comparação ao BMH (Hipótese)
  - Dependerá do tamanho alfabeto/texto que está sendo trabalhado
- Apresenta mais eficiência em textos de linguagens naturais

## Questão 5)

### A)

- E QUEM AMAFAGAFAR OS MAFAGAFINHOS BOM AMAFAGAFINHADOR SERÁ AMAFA
  - Colisão entre A do Padrão e E do Texto
- E QUEM **AMAFAGAFAR** OS MAFAGAFINHOS BOM AMAFAGAFINHADOR SERÁ AMAFA
  - Colisão entre o F do padrão e M do texto
- E QUEM **AMAFAGAFAR** OS MAFAGAFINHOS BOM AMAFAGAFINHADOR SERÁ **AMAFA**
  - Padrão encontrado no texto

### B)

- Tabela de Deslocamentos:

$$\begin{aligned}d[a] &= 1 \\d[m] &= 4 \\d[f] &= 2\end{aligned}$$

- E QUEM**AMAFAGAFAR** OS MAFAGAFINHOS BOM AMAFAGAFINHADOR SERÁ AMAFA
  - Colisão entre A do padrão e E do texto , deslocamos 4 unidades (M)
- E QUEM AMAFAGAFAR OS MAFAGAFINHOS BOM AMAFAGAFINHADOR SERÁ AMAFA
  - Colisão entre o A do padrão e M do texto, deslocamos 1 unidade (A)
- E QUEM AMAFAGAFAR OS MAFAGAFINHOS BOM AMAFAGAFINHADOR SERÁ AMAFA
  - Colisão entre o F do padrão e M do texto, deslocamos 2 unidades (F)
- E QUEM **AMAFAGAFAR** OS MAFAGAFINHOS BOM AMAFAGAFINHADOR SERÁ **AMAFA**
  - Padrão encontrado no texto