Inhalt

[4.10 1](#_Toc29477210)

[4.10.1 1](#_Toc29477211)

[4.10.2 3](#_Toc29477212)

[4.11 3](#_Toc29477213)

[4.13 4](#_Toc29477214)

[4.13.1 4](#_Toc29477215)

[4.13.2 4](#_Toc29477216)

[4.13.3 5](#_Toc29477217)

[4.14 5](#_Toc29477218)

[4.14.1 5](#_Toc29477219)

[4.14.2 5](#_Toc29477220)

## 4.10

### 4.10.1

ABRACADABRA

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| j | pat.charAt(j) | dfa[][j] | | | | |  |  |
|  |  | A | B | C | D | R | ABRACADABRA | |
| 0 | A | 1 |  |  |  |  | **A** |  |
|  |  |  |  |  |  |  | B |  |
|  |  |  | 0 |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  |  | C |  |
|  |  |  |  | 0 |  |  |  | ABRACADABRA |
|  |  |  |  |  |  |  | D |  |
|  |  |  |  |  | 0 |  |  | ABRACADABRA |
|  |  |  |  |  |  |  | R |  |
|  |  |  |  |  |  | 0 |  | ABRACADABRA |
| 1 | B |  | 2 |  |  |  | **AB** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  | 1 |  |  |  |  | AA |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | AC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | AD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | AR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 2 | R |  |  |  |  | 3 | **ABR** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  | 1 |  |  |  |  | ABA |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 0 |  |  |  | ABB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 3 | A | 4 |  |  |  |  | **ABRA** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 0 |  |  |  | ABRB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABRC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 4 | C | 5 |  |  |  |  | **ABRAC** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 1 |  |  |  | ABRAA |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABRAB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRAC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRAR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 5 | A | 7 |  |  |  |  | **ABRACA** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 0 |  |  |  | ABRACB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABRACC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRACD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRACR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 6 | D | 7 |  |  |  |  | **ABRACAD** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 1 |  |  |  | ABRACAA |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 2 |  |  | ABRACAB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRACAC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRACAR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 7 | A | 7 |  |  |  |  | **ABRACADA** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 0 |  |  |  | ABRACADB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABRACADC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRACADD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRACADR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 8 | A | 7 |  |  |  |  | **ABRACADAB** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 1 |  |  |  | ABRACADAA |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABRACADAC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRACADAD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRACADAR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 9 | A | 7 |  |  |  |  | **ABRACADABR** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 1 |  |  |  | ABRACADABA |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABRACADABB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRACADABC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRACADABD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
| 10 | A | 7 |  |  |  |  | **ABRACADABRA** |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  | 0 |  |  |  | ABRACADABRB |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  | 0 |  |  | ABRACADABRC |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  | 0 |  | ABRACADABRD |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |
|  |  |  |  |  |  | 0 | ABRACADABRR |  |
|  |  |  |  |  |  |  |  | ABRACADABRA |

### 4.10.2

To implement the mismatched character heuristic, we use an array right[] that gives, for each character in the alphabet, the index of its rightmost occurrence in the pattern (or -1 if the character is not in the pattern).

ABRACADABRA

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A** | **B** | **R** | **A** | **C** | **A** | **D** | **A** | **B** | **R** | **A** |  |
| **c** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **right[c]** |
| **A** |  |  |  |  |  |  |  |  |  |  | 10 | **10** |
| **B** |  | 1 |  |  |  |  |  |  |  |  |  | **1** |
| **C** |  |  |  |  | 4 |  |  |  |  |  |  | **4** |
| **D** |  |  |  |  |  |  | 6 |  |  |  |  | **6** |
| **R** |  |  |  |  |  |  |  |  |  | 9 |  | **9** |

## 4.11

Boyer-Moore

TXTXXTXUTXTXUXTXXUTTTTTTXUUTXTTT

TXUUT

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| T | X | T | X | X | T | X | U | T | X | T | X | U | X | T | X | X | U | T | T | T | T | T | T | X | U | U | T | X | T | T | T |
| T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | T | X | U | U | T |  |  |  |  |

Rabin-Karp

hash(TXUUT) = h0

TUXTXXTXUTXTXUXTXXUTTTTTTXUUTXTTT

|  |  |  |
| --- | --- | --- |
| | x | | hash(x) | == h0 ? |
| | TUXTX | XTXUTXTXUXTXXUTTTTTTXUUTXTTT | h1 | no |
| T | UXTXX | TXUTXTXUXTXXUTTTTTTXUUTXTTT | h2 | no |
| TU | XTXXT | XUTXTXUXTXXUTTTTTTXUUTXTTT | h3 | no |
| … |  |  |
| TUXTXXTXUTXTXUXTXXUTTTTT | TXUUT | XTTT | h0 | yes |
| T == T  X == X  U == U  U == U  T == T  → match | | |
| … |  |  |
| TUXTXXTXUTXTXUXTXXUTTTTTTXUU | T XTTT | | h4 | no → done |

## 4.13

### 4.13.1

Der eine Buchstabe würde durch 0 und der andere durch 1 repräsentiert werden

### 4.13.2

RunLength: 1a1b, 1a1b1a1b, …

LZW: 81, 8181, 81818181

|  |  |
| --- | --- |
| codeword table | |
| key | value |
| ab | 81 |

### 4.13.3

Lauflänge: ?N

Kompressionsrate für kleine Binärdatei 32/40=80%

Kompressionsrate für Textdateien 416/96=433% KEINE LAUFLÄNGENKODIERUNG FÜR ASCIII!

Bitmap: Kompressionsrate 1144/1536= 74%

Höher aufgelöste Bitmap: Kompressionsrate 2296/6144 = 37%

Hufmann: 0,53N

Bei mittelgroßen Texten erreicht Huffman eine Kompressionsrate von ≈ 53%

LZW: 0,46N

Komprimierungsrate 2667936 / 5812560 = 46 % bisher das beste Ergebnis

## 4.14

### 4.14.1

Es kommt auf die Sortierung nach dem nach Häufigkeit sortiert wurde an. Je nachdem kann die Kodierung abweichen

### 4.14.2

Der am schnellsten zu erreichende Knoten und somit der mit der kürzesten Bitfolge ist das häufigste und der am langsamsten zu erreichende Knoten, der unterste, und somit der mit der längsten Bitfolge ist das am wenigsten häufige Zeichen

4.15

EINS+EINSISTNICHTEINS

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| + | C | E | I | N | H | S | T | EOF |
| 2B | 43 | 45 | 49 | 4E | 48 | 53 | 54 | 80 |

EI 81

IN 82

NS 83

S+ 84

+E 85

EIN 86

NSI 87

IS 88

ST 89

TN 8A

NI 8B

IC 8C

CH 8D

HT 8E

TE 8F

EINS 90

SEOF 91

----------------

45 49 4E 53 2B 86 87 49 53 54 4E 49 43 48 54 90 53 80