Assignment 7

Instructions:

Working modulo q=11, how many spurious hits does the Rabin-Karp matcher encounter in the text T = 314159265358993 when looking for the pattern P = 26?

Show the working of Rabin-Karp algorithm step by step and graphically indicate spurious hits while showing each iteration.

Solution :

**RABIN-KARP-MATCHER (T, P, d, q):**

n = T.length

m = P.length

h = dm-1 mod q

p = 0

t0 = 0

for i ← 1 to m //preprocessing

do p = (dp + P[i]) mod q

t0 = (dt0+T [i]) mod q

for s ← 0 to n-m //matching

if p == ts

if P [1..m] == T [s+1..s + m]

print "Pattern occurs with shift" s

if s < n-m

ts+1 = (d (ts-T [s+1]h) + T [s+m+1]) mod q

**Working of Rabin-Karp algorithm on,**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T=

|  |  |
| --- | --- |
| 2 | 6 |

Pattern P =

n = T.length = 15

m = P.length = 2

Here, working modulo q = 11 and

p = P mod q = 26 mod 11 = 4

Let **ts** denotes modulo **q** for text of length **m = 2**

Now let’s find the exact match of **P mod q**

For s = 0

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t0** = 31 mod 11 = 9

For s = 1

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t1** = 14 mod 11 = 3

For s = 2

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t2** = 41 mod 11 = 8

For s = 3

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t3** = 15 mod 11 = 4 //Spurious Hit

For s = 4

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t4** = 59 mod 11 = 4 //Spurious Hit

For s = 5

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t5** = 92 mod 11 = 4 //Spurious Hit

For s = 6

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t6** = 26 mod 11 = 4 //Exact Match

**Pattern occurs with shift 6**

For s = 7

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t7** = 65 mod 11 = 10

For s = 8

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t8** = 53 mod 11 = 9

For s = 9

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t0** = 35 mod 11 = 2

For s = 10

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t10** = 58 mod 11 = 3

For s = 11

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t11** = 89 mod 11 = 1

For s = 12

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t12** = 99 mod 11 = 0

For s = 13

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

**t13** = 93 mod 11 = 5

**Answer :**

q = 11

|  |  |
| --- | --- |
| 2 | 6 |

Pattern = **p = 4**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 1 | 4 | 1 | 5 | 9 | 2 | 6 | 5 | 3 | 5 | 8 | 9 | 9 | 3 |

Text T =

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 9 | 3 | 8 | 4 | 4 | 4 | 4 | 10 | 9 | 2 | 3 | 1 | 0 | 5 |

**ts** =

Valid Match

Spurious Hits

So, the **Rabin-Karp** matcher encounters **3 spurious hits** and exactly one valid match with shift 6.