Naïve String Matching

$P[1..m] = T[s+1..s+m]$ for each of the $n-m+1$ possible values of s.

NAIVE-STRING-MATCHER(T, P)

```
1  n = T.length
2  m = P.length
3  for s = 0 to n − m
4      if P[1..m] == T[s + 1..s + m]
5          print "Pattern occurs with shift" s
```

$$O((n-m+1)m)$$

The Rabin-Karp algorithm

$P[1..m]$ let p denote its corresponding decimal value.

$T[1..n]$ let t_s denote the decimal value

$T[s+1..s+m]$, for $s = 0, 1, \ldots, n-m$

$t_s = p$ $\qquad$ $T[s+1..s+m] = P[1..m]$

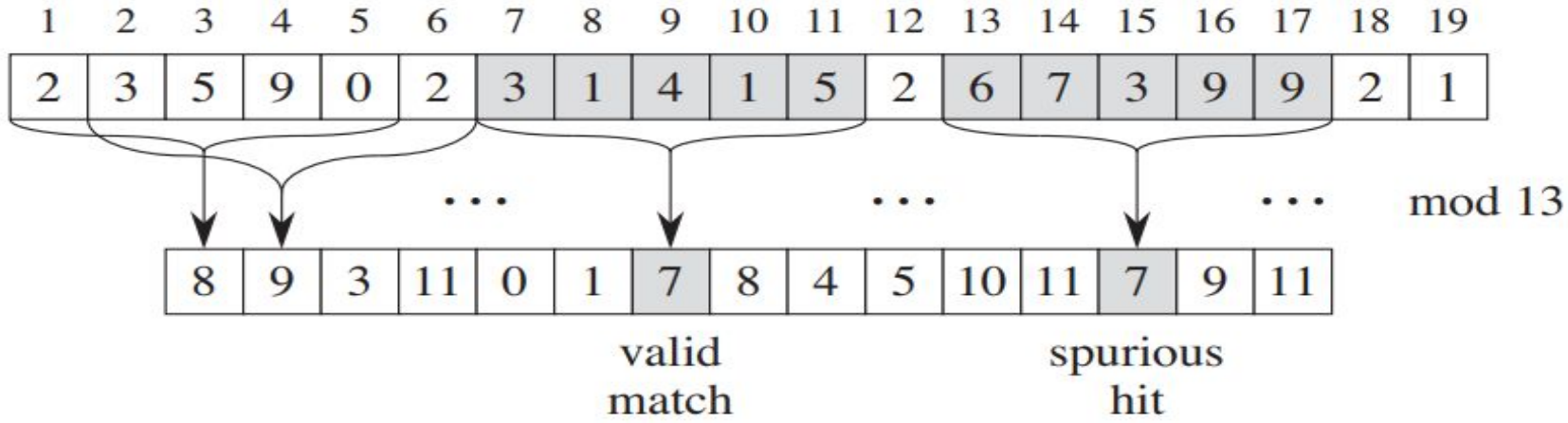
The Rabin-Karp algorithm

$$t_{s+1} = 10(t_s - 10^{m-1}T[s+1]) + T[s+m+1]$$

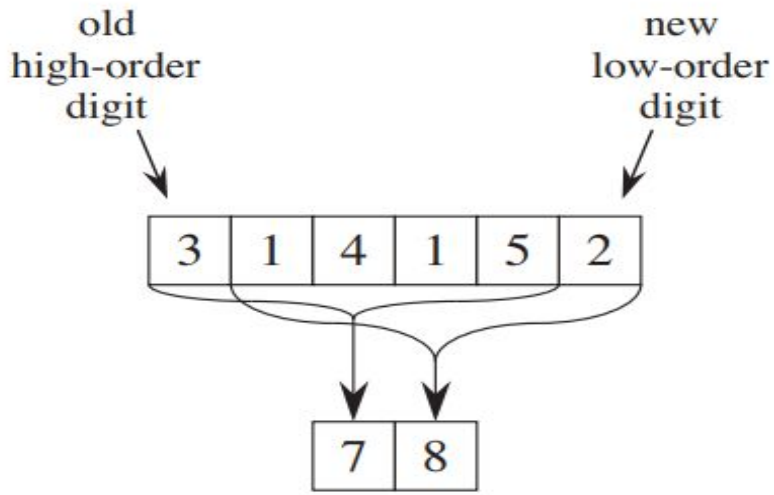
$$t_s = 31415.$$

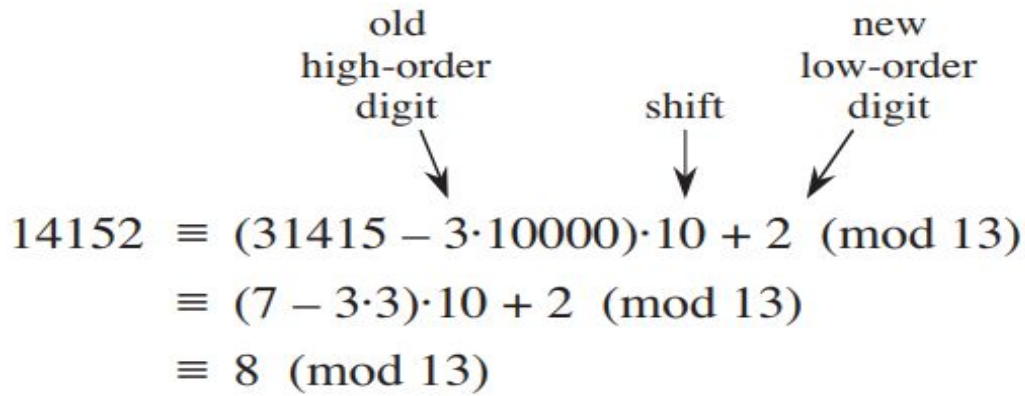
$$\begin{aligned} t_{s+1} &= 10(31415 - 10000 \cdot 3) + 2 \\ &= 14152 . \end{aligned}$$

$$t_{s+1} = (d(t_s - T[s+1]h) + T[s+m+1]) \bmod q$$


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
2 3 5 9 0 2 3 1 4 1 5 2 6 7 3 9 9 2 1
mod 13
8 9 3 11 0 1 7 8 4 5 10 11 7 9 11
valid match
spurious hit

(b)


old high-order digit
new low-order digit
3 1 4 1 5 2
7 8


old high-order digit
shift
new low-order digit
14152 ≡ (31415 – 3·10000)·10 + 2 (mod 13)
≡ (7 – 3·3)·10 + 2 (mod 13)
≡ 8 (mod 13)

(c)

```
RABIN-KARP-MATCHER(T, P, d, q)
 1  n = T.length
 2  m = P.length
 3  h = d^(m−1) mod q
 4  p = 0
 5  t_0 = 0
 6  for i = 1 to m                          // preprocessing
 7      p = (dp + P[i]) mod q
 8      t_0 = (dt_0 + T[i]) mod q
 9  for s = 0 to n − m                      // matching
10      if p == t_s
11          if P[1..m] == T[s + 1..s + m]
12              print "Pattern occurs with shift" s
13      if s < n − m
14          t_{s+1} = (d(t_s − T[s + 1]h) + T[s + m + 1]) mod q
```

The Rabin-Karp algorithm

$\Theta((n - m + 1)m)$