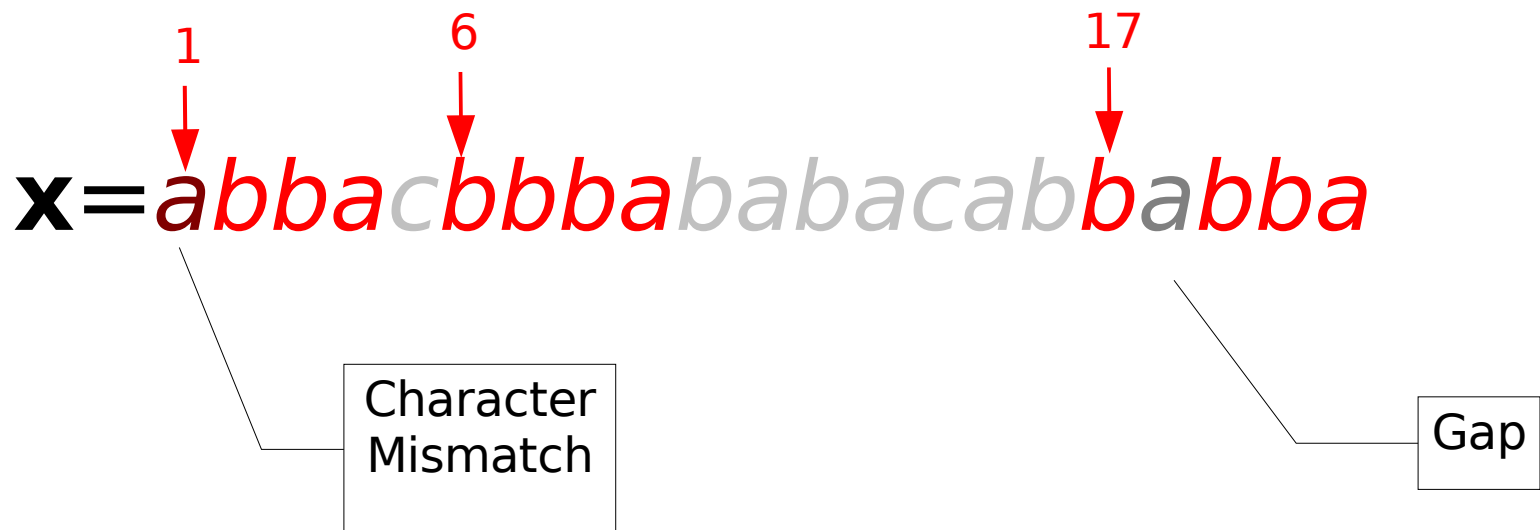


# Approximate k-edit-distance

---

Approximative ( $k$ -distance) matching, now for edit distance

Given string  $x = \textit{abbacbbbabacabbba}$  and pattern  $p = \textit{bbba}$  find all “almost”-occurrences of  $p$  in  $x$



# Edit distance

---

The *edit-distance* between strings  $x$  and  $y$  is the minimal number of

- insertions
- deletions
- substitutions

needed to translate  $x$  into  $y$

$d(abab, acc) = 3$ :  $abab \rightarrow aba \rightarrow aca \rightarrow acc$

$d(abab, aac) = 2$ :  $abab \rightarrow aab \rightarrow aac$

# Calculating the edit-distance

---

Basis cases:

- string vs empty string:

$$d(\mathbf{x}, "") = d("", \mathbf{x}) = |\mathbf{x}|$$

- two single characters:

$$d(a, b) = \begin{cases} 1 & \text{if } a \neq b \\ 0 & \text{if } a = b \end{cases}$$

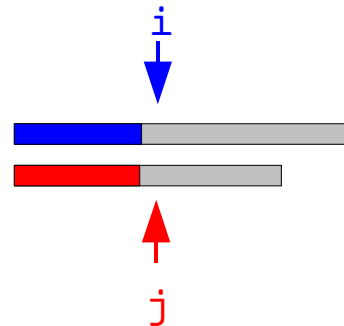
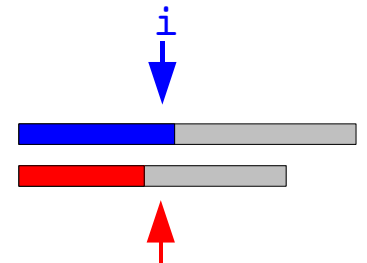
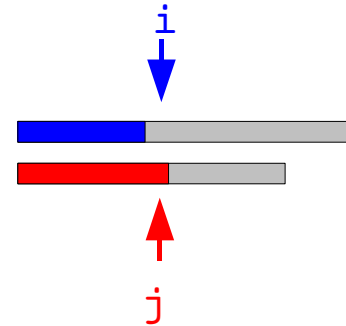
# Calculating the edit-distance

Recursion:

- two non-empty strings:

$$d(\mathbf{x}[1..i], \mathbf{y}[1..j]) =$$

$$\min \begin{cases} d(\mathbf{x}[1..i-1], \mathbf{y}[1..j]) + 1 \\ d(\mathbf{x}[1..i], \mathbf{y}[1..j-1]) + 1 \\ d(\mathbf{x}[1..i-1], \mathbf{y}[1..j-1]) + d(\mathbf{x}[i], \mathbf{y}[j]) \end{cases}$$



# Dynamic programming algorithm

Use table  $c[i,j] = d(x[1..i],y[1..j])$

Initialize:

$c[0,0] = 0$

**for**  $i=1..|x|$ :  $c[i,0] = i$

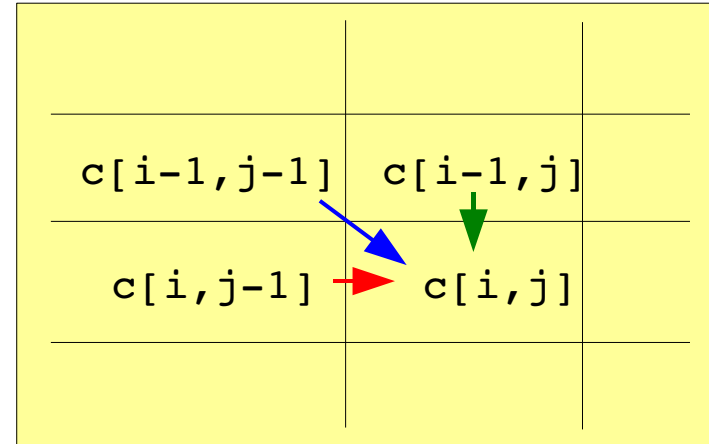
**for**  $j=1..|p|$ :  $c[0,j] = j$

Main algorithm:

**for**  $i=1..|x|$ :

**for**  $j=1..|p|$ :

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$



# Example

---

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

**a b a a c b c c**

	0	1	2	3	4	5	6	7	8
<b>a</b>	1								
<b>b</b>	2								
<b>a</b>	3								
<b>b</b>	4								
<b>c</b>	5								
<b>a</b>	6								

# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

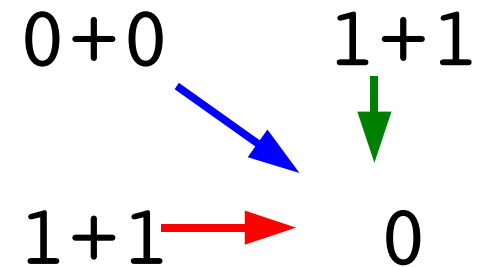
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

	a	b	a	a	c	b	c	c
0	1	2	3	4	5	6	7	8
a	1	<u>0</u>						
b	2							
a	3							
b	4							
c	5							
a	6							



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

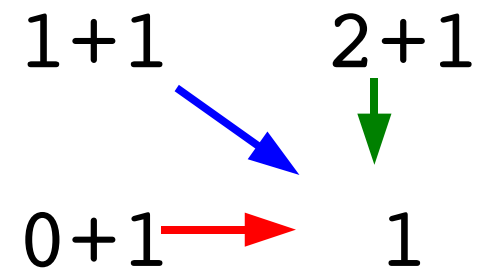
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	<u>1</u>						
b	2								
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

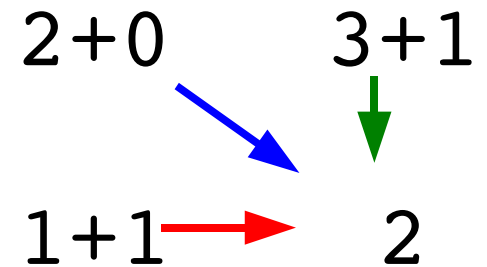
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

	a	b	a	a	c	b	c	c
0	1	2	3	4	5	6	7	8
a	1	0	1	<u>2</u>				
b	2							
a	3							
b	4							
c	5							
a	6							



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

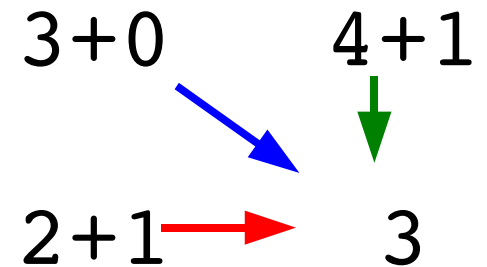
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	<u>3</u>				
b	2								
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

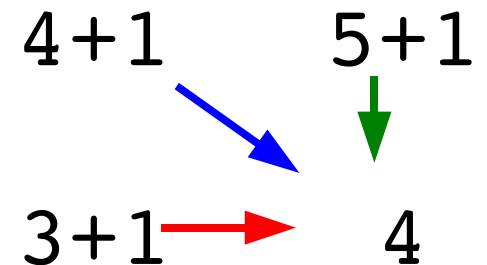
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

	a	b	a	a	c	b	c	c
0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	<u>4</u>		
b	2							
a	3							
b	4							
c	5							
a	6							



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

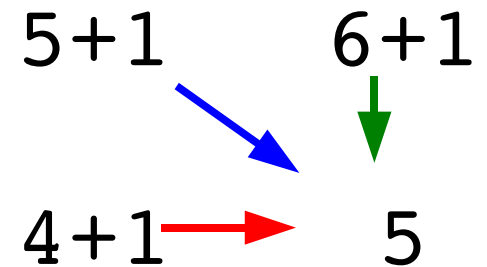
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	<u>5</u>		
b	2								
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

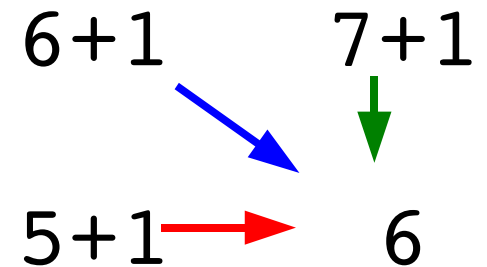
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	<u>6</u>	
b	2								
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

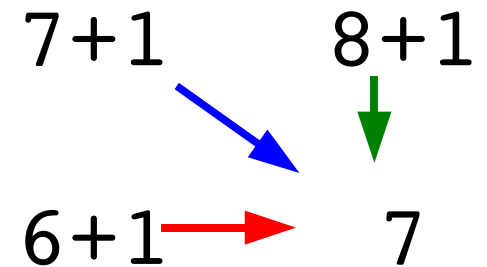
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

	a	b	a	a	c	b	c	c	
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	<u>7</u>
b	2								
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

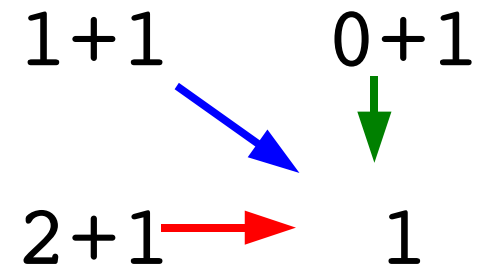
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	<u>1</u>							
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

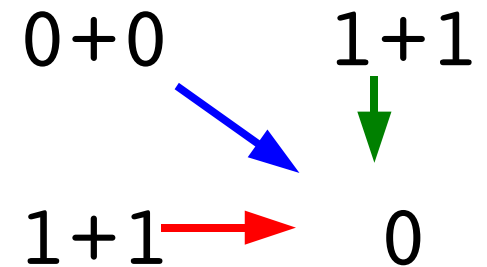
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	<u>0</u>						
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

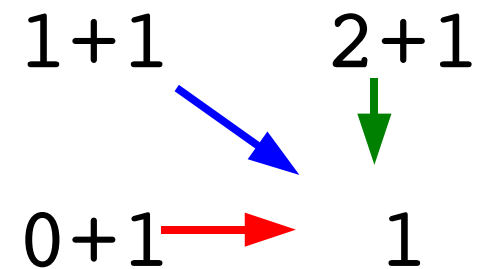
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	0	<u>1</u>					
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

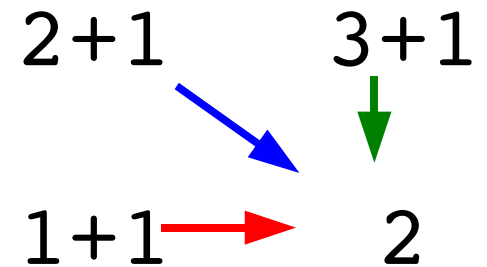
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	0	1	<u>2</u>				
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

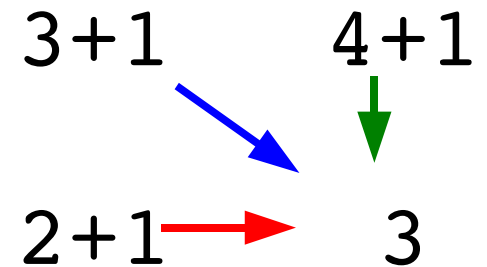
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	0	1	2	<u>3</u>			
a	3								
b	4								
c	5								
a	6								



# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

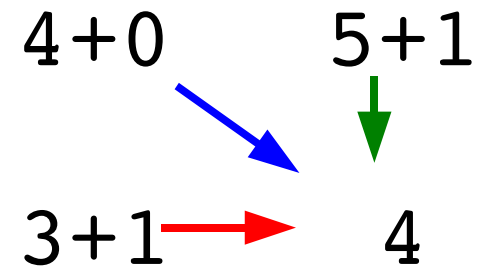
```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

	a	b	a	a	c	b	c	c	
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	0	1	2	3	<u>4</u>		
a	3								
b	4								
c	5								
a	6								





# Example

```
c[0,0] = 0
```

```
for i=1..|x|: c[i,0] = i
```

```
for j=1..|p|: c[0,j] = j
```

```
for i=1..|x|:
```

```
  for j=1..|p|:
```

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(x[i],y[j]) \end{cases}$$

		a	b	a	a	c	b	c	c
	0	1	2	3	4	5	6	7	8
a	1	0	1	2	3	4	5	6	7
b	2	1	0	1	2	3	4	5	6
a	3	2	1	0	1	2	3	4	5
b	4	3	2	1	1	2	2	3	4
c	5	4	3	2	2	1	2	2	3
a	6	5	4	3	2	2	2	3	3

# Dynamic programming algorithm

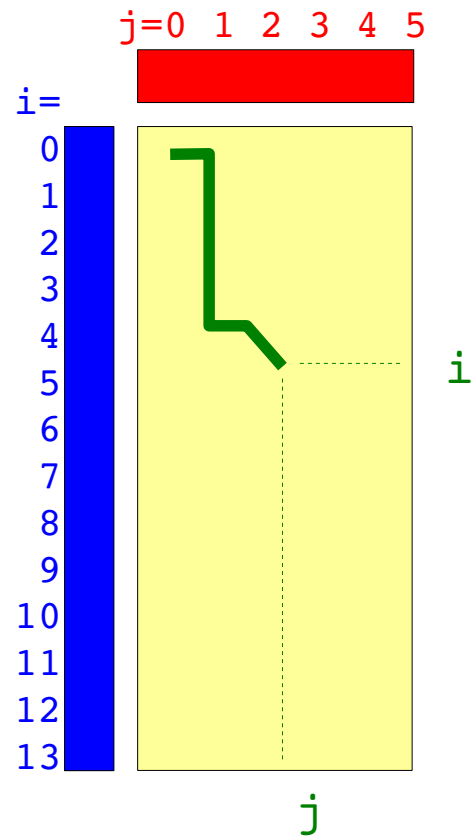
After filling out  $c$ ,  $d(\mathbf{x}, \mathbf{y}) = c[|\mathbf{x}|, |\mathbf{y}|]$

Time and space complexity:  $O(|\mathbf{x}||\mathbf{y}|)$

# Approximate pattern matching

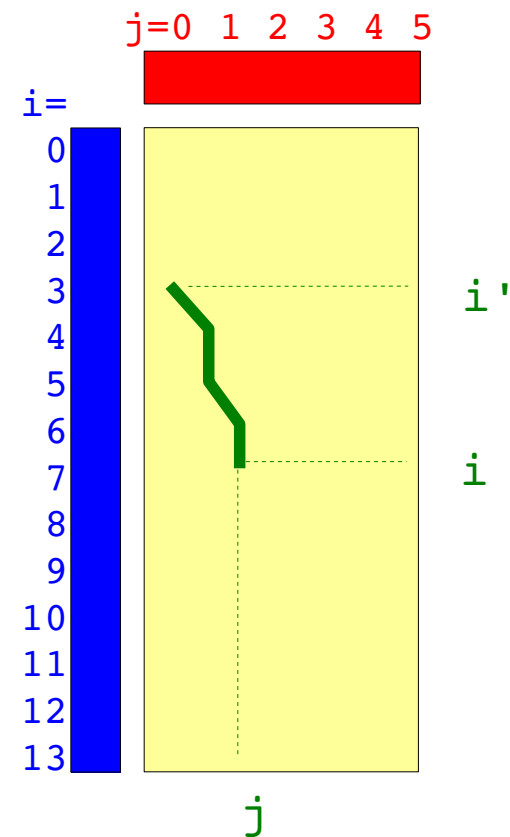
Edit distance

$$c[i,j] = d(\mathbf{x}[1..i], \mathbf{y}[1..j])$$



Edit distance  
pattern matching

$$c[i,j] = \min_{i' \leq i} d(\mathbf{x}[i'..i], \mathbf{y}[1..j])$$



# Approximate pattern matching

Use table  $c[i,j] = \min_{i' \leq i} d(\mathbf{x}[1..i], \mathbf{y}[1..j])$

Initialize:

$c[0,0] = 0$

**for**  $i=1..|\mathbf{x}|$ :  $c[i,0] = \underline{0}$

**for**  $j=1..|\mathbf{p}|$ :  $c[0,j] = j$

Makes it free to start  
at **any** index  $i$ '

Main algorithm:

**for**  $i=1..|\mathbf{x}|$ :

**for**  $j=1..|\mathbf{p}|$ :

$$c[i,j] = \min \begin{cases} c[i-1,j] + 1 \\ c[i,j-1] + 1 \\ c[i-1,j-1] + d(\mathbf{x}[i], \mathbf{y}[j]) \end{cases}$$

# Approximate pattern matching

After filling out, in time and space  $O(|x||p|)$  all indices  $i$ , where  $c[i, |p|] \leq k$ , correspond to one or more approximate matches.

Some backtracking is needed to find the corresponding  $i'$  indices ... we can find one  $i'$  for each  $i$  in time  $O(|p|)$  per  $i$ , for a total time of  $O(|x||p|)$ .  
(Theorem 10.1.1)

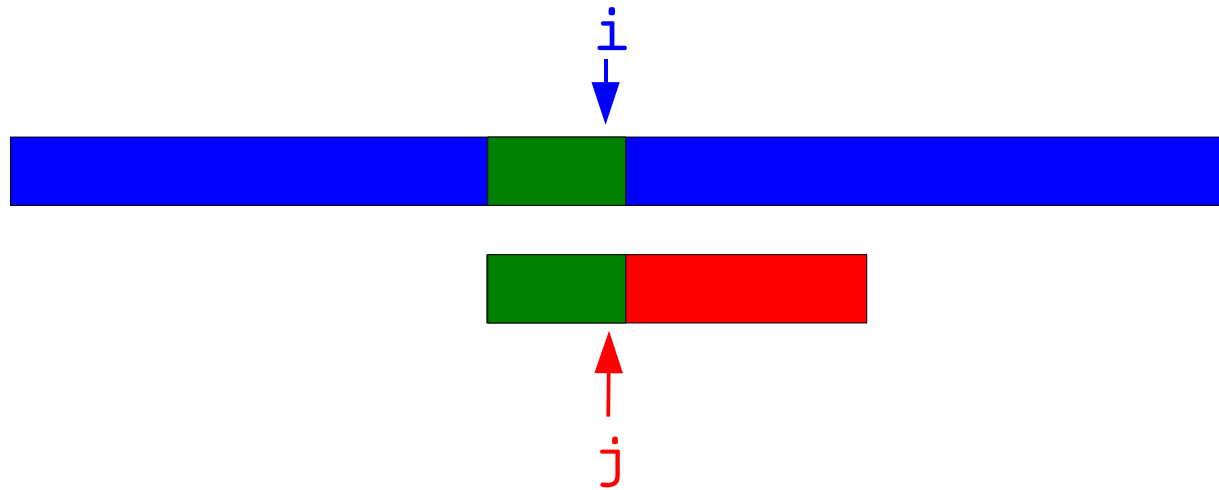
(More complicated to get all matches...)

# Wu-Manber

---

We define a *matrix*  $s$  – the state of matching so far – by:

$s[q,j] = 0$  iff  $d(x[i-j+1 .. i], p[1..j]) \leq q$   
for  $j=0..|p|$ , and  $q=0..k$



# Wu-Manber

---

As before, we use a pre-calculated bit-matrix:

$$\mathbf{t}[h, j] = \begin{cases} 0 & \text{if } \mathbf{p}[j] == h \\ 1 & \text{if } \mathbf{p}[j] != h \end{cases}$$

with rows indexed by the alphabet and  
columns indexed by indices in  $\mathbf{p}$

# Wu-Manber

---

The recursion:

$$c[i, j] = \min \begin{cases} c[i-1, j] + 1 \\ c[i, j-1] + 1 \\ c[i-1, j-1] + d(x[i], y[j]) \end{cases}$$

becomes:

$$s^i[q, j] = \& \begin{cases} s^{i-1}[q-1, j] \\ s^i[q-1, j-1] \\ s^{i-1}[q-1, j-1] \& (s^{i-1}[q, j-1] | t[x[i], j]) \end{cases}$$

# Wu-Manber

---

The expression:

$$s^i[q, j] = \& \left\{ \begin{array}{l} s^{i-1}[q-1, j] \\ s^i[q-1, j-1] \\ s^{i-1}[q-1, j-1] \& (s^{i-1}[q, j-1] \mid t[x[i], j]) \end{array} \right.$$

can be computed as:

```
old = s
```

```
s[0] = (old[0] >> 1) | t[x[i]] // SHIFT-and-OR
```

```
for q=1..k:
```

```
  s1 = old[q-1] // s1[j] = s^{i-1}[q-1, j]
```

```
  s2 = s[q-1] >> 1 // s2[j] = s^i[q-1, j-1]
```

```
  s3 = s1 >> 1 // s3[j] = s^{i-1}[q-1, j-1]
```

```
  s4 = old[q] >> 1 // s4[j] = s^{i-1}[q, j-1]
```

```
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

# Wu-Manber

---

Special case:

-Initial matrix:  $s[q] = 01^{|p|}$

Match when  $s[k, |p|] == 0$

# Example

---

$i=0$



**x** = *bbacbbbababacabbba*

**p** = *bbba*

**01234**

**s**<sup>0</sup>[0]: 01111

**s**<sup>0</sup>[1]: 01111

**s**<sup>0</sup>[2]: 01111

# Example

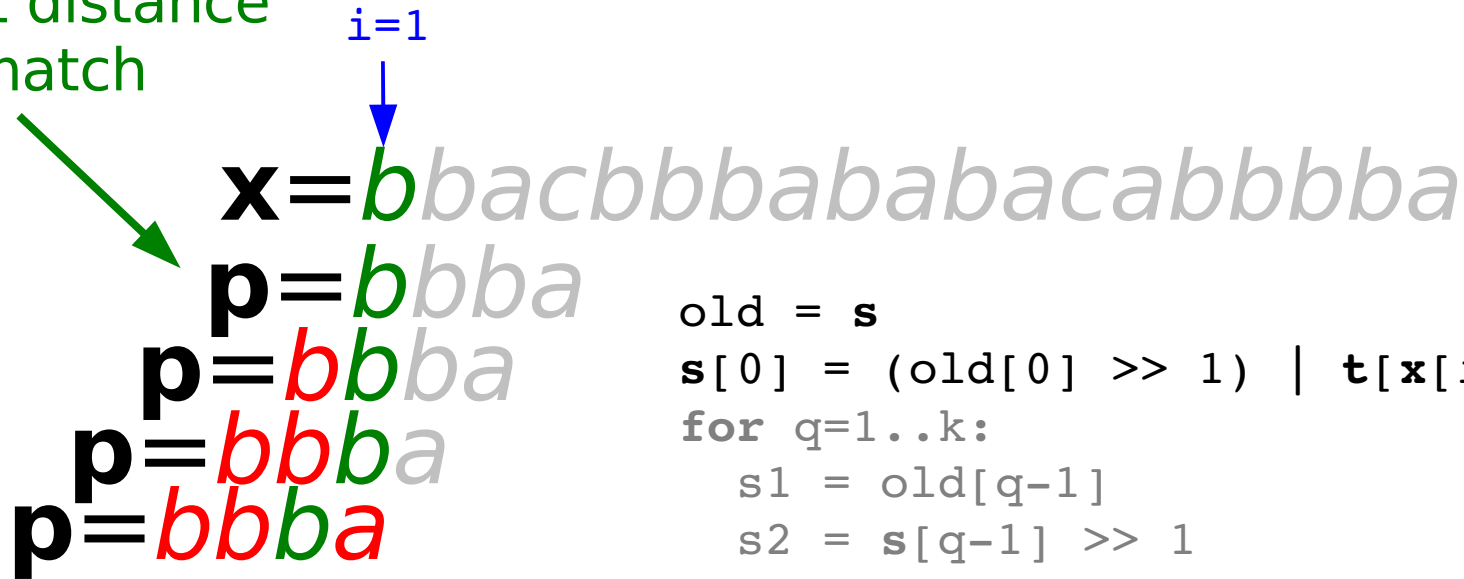
$i=1$   
↓  
 $x = b$ *bacbbbababacabbba*  
 $p = b$ *bba*  
 $p = b$ *bbba*  
 $p = b$ *bbbba*  
 $p = b$ *bbbba*

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>
$s^0[0]:$	01111	$s^1[0]:$	00111 = 00111   00001
$s^0[1]:$	01111	$s^1[1]:$	
$s^0[2]:$	01111	$s^1[2]:$	

# Example

0 edit distance  
match

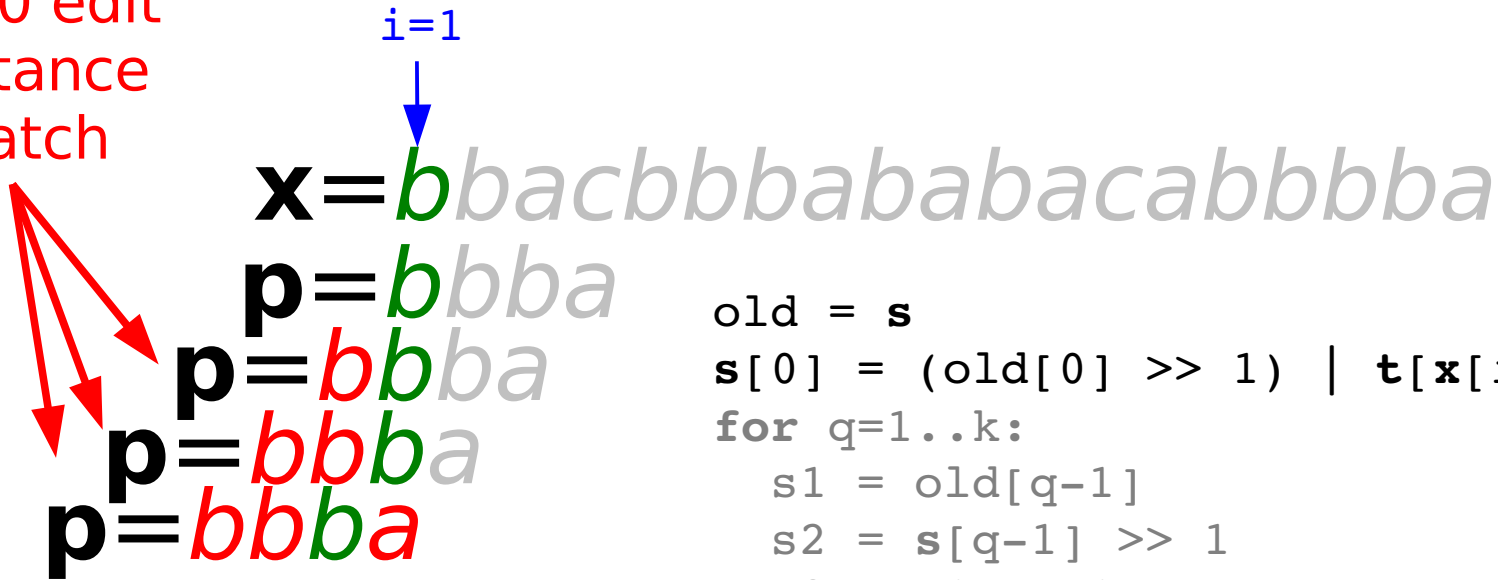


```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	01234		01234
$s^0[0]:$	01111	$s^1[0]:$	<u>0</u> 0111 = 00111   00001
$s^0[1]:$	01111	$s^1[1]:$	
$s^0[2]:$	01111	$s^1[2]:$	

# Example


Not 0 edit  
distance  
match



```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>
$s^0[0]:$	01111	$s^1[0]:$	00 <u>111</u> = 00111   00001
$s^0[1]:$	01111	$s^1[1]:$	
$s^0[2]:$	01111	$s^1[2]:$	

# Example

$i=1$   
  
 $x = b$ *bacbbbababacabbba*  
 $p = b$ *bba*  
 $p = b$ *ba*  
 $p = b$ *ba*  
 $p = b$ *ba*

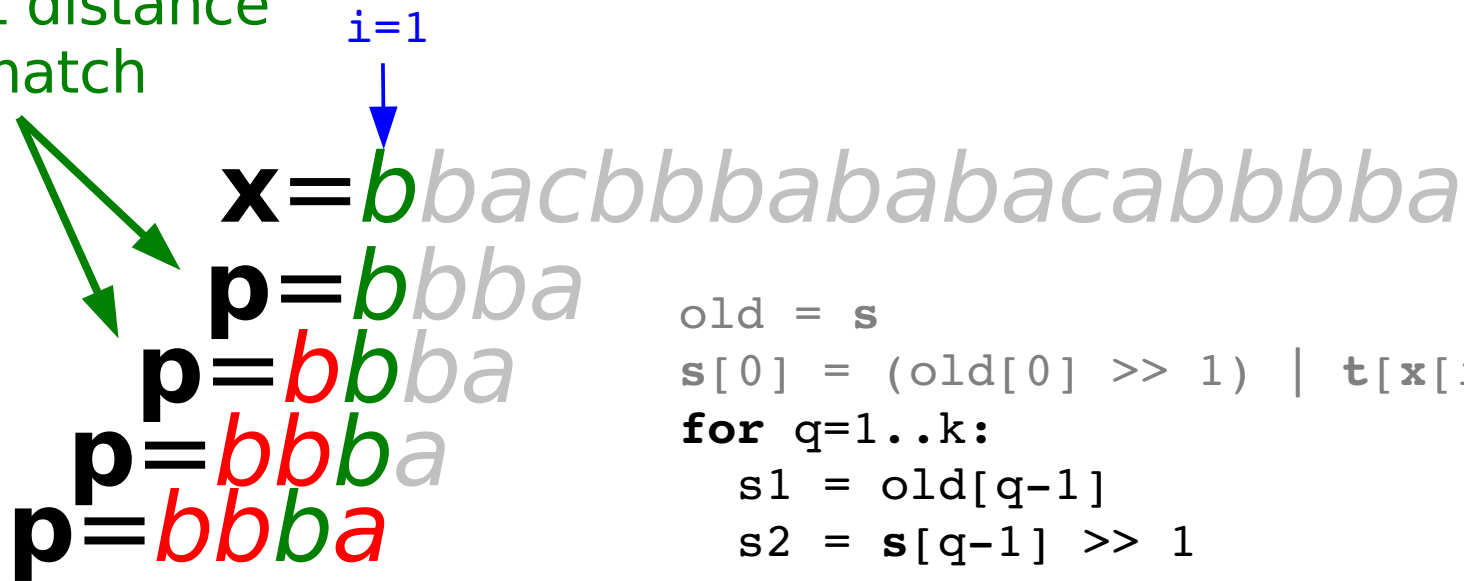
```

old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
  
```

	01234		01234
$s^0[0]:$	01111	$s^1[0]:$	00111
$s^0[1]:$	01111	$s^1[1]:$	00011 = 01111 & 00011
$s^0[2]:$	01111	$s^1[2]:$	& 00111 & (00111   00001)

# Example

1 edit distance  
match

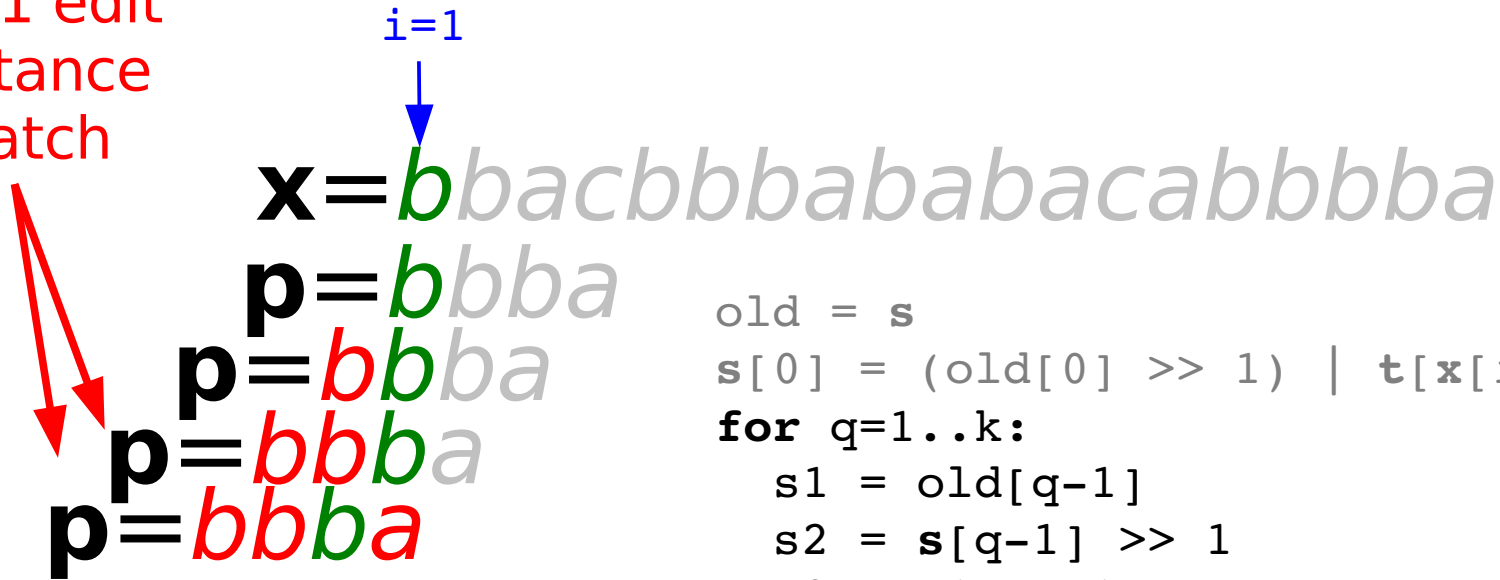


```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
    s1 = old[q-1]
    s2 = s[q-1] >> 1
    s3 = s1 >> 1
    s4 = old[q] >> 1
    s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
$s^0[0]:$	01111	$s^1[0]:$	00111	
$s^0[1]:$	01111	$s^1[1]:$	00011 = 01111 & 00011	
$s^0[2]:$	01111	$s^1[2]:$	& 00111 & (00111   00001)	

# Example


Not 1 edit  
distance  
match



```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	01234		01234
$s^0[0]:$	01111	$s^1[0]:$	00111
$s^0[1]:$	01111	$s^1[1]:$	000 <u>11</u> = 01111 & 00011
$s^0[2]:$	01111	$s^1[2]:$	& 00111 & (00111   00001)

# Example

$i=1$   
  
 $x = b$ *bacbbbababacabbba*  
 $p = b$ *bba*  
 $p = b$ *ba*  
 $p = b$ *ba*  
 $p = b$ *ba*

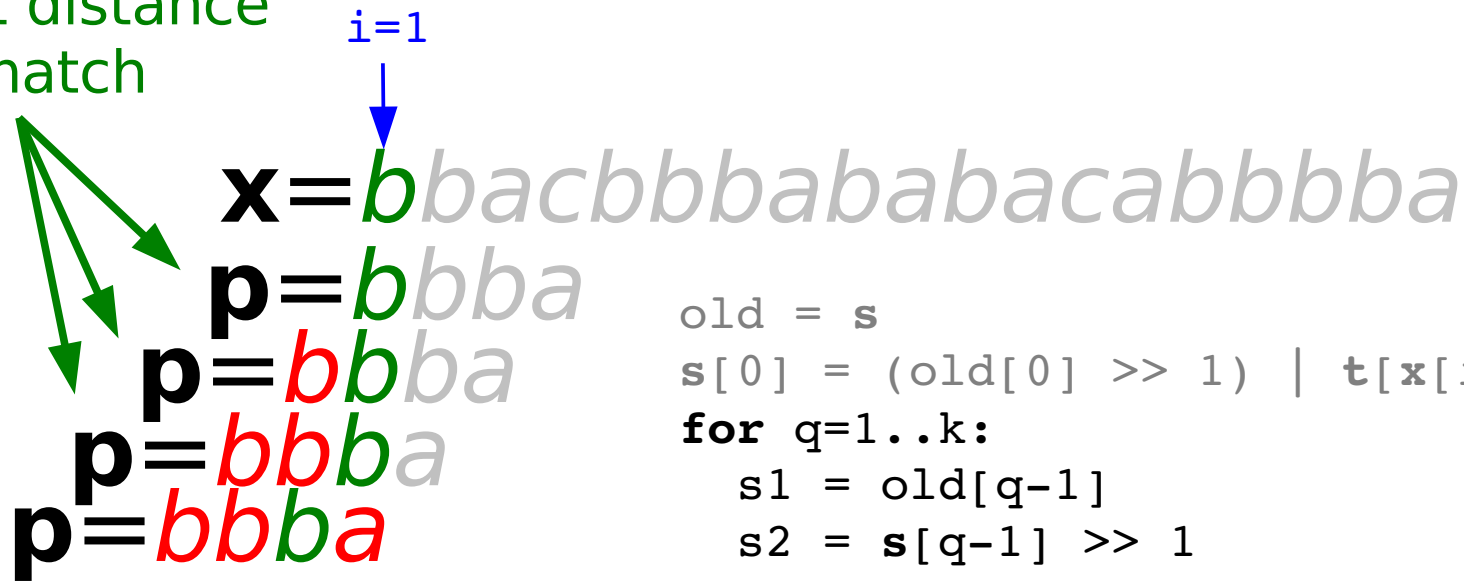
```

old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
  
```

	<b>01234</b>		<b>01234</b>	
$s^0[0]:$	01111	$s^1[0]:$	00111	
$s^0[1]:$	01111	$s^1[1]:$	00011	
$s^0[2]:$	01111	$s^1[2]:$	00001	$= 01111 \ \& \ 00001$
				$\& \ 00111 \ \& \ (00111 \   \ 00001)$

# Example

2 edit distance  
match

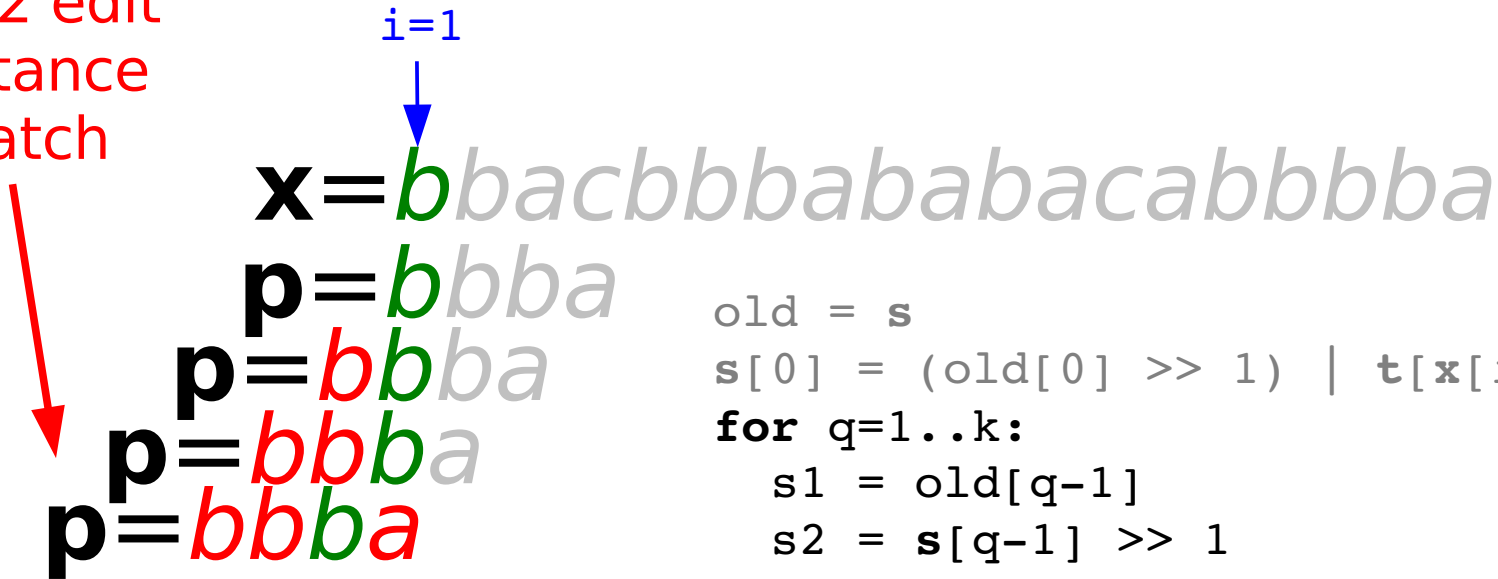


```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
    s1 = old[q-1]
    s2 = s[q-1] >> 1
    s3 = s1 >> 1
    s4 = old[q] >> 1
    s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
$s^0[0]:$	01111	$s^1[0]:$	00111	
$s^0[1]:$	01111	$s^1[1]:$	00011	
$s^0[2]:$	01111	$s^1[2]:$	<u>00001</u>	$= 01111 \ \& \ 00001$
				$\& \ 00111 \ \& \ (00111 \   \ 00001)$

# Example

Not 2 edit  
distance  
match



```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
$s^0[0]:$	01111	$s^1[0]:$	00111	
$s^0[1]:$	01111	$s^1[1]:$	00011	
$s^0[2]:$	01111	$s^1[2]:$	0000 <u>1</u>	$= 01111 \ \& \ 00001$
				$\& \ 00111 \ \& \ (00111 \   \ 00001)$

# Example

$i=2$   
↓

$x = bbacbbbababacabbba$

$p = bbba$

$p = bbba$

$p = bbba$

$p = bbba$

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	01234		01234
$s^1[0]:$	00111	$s^2[0]:$	00011 = 00011   00001
$s^1[1]:$	00011	$s^2[1]:$	
$s^1[2]:$	00001	$s^2[2]:$	

# Example

0 edit distance  
match

$i=2$

$x = bbacbbbababacabbba$

$p = bbba$

$p = bbba$

$p = bbba$

$p = bbba$

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
$s^1[0]:$	00111	$s^2[0]:$	<u>000</u> 11	= 00011   00001
$s^1[1]:$	00011	$s^2[1]:$		
$s^1[2]:$	00001	$s^2[2]:$		

# Example

Not 0 edit  
distance  
match

$i=2$   
 $x = bbacbbbababacabbba$   
 $p = bbba$   
 $p = bbba$   
 $p = bbba$   
 $p = bbba$

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
$s^1[0]:$	00111	$s^2[0]:$	000 <u>11</u>	= 00011   00001
$s^1[1]:$	00011	$s^2[1]:$		
$s^1[2]:$	00001	$s^2[2]:$		

# Example

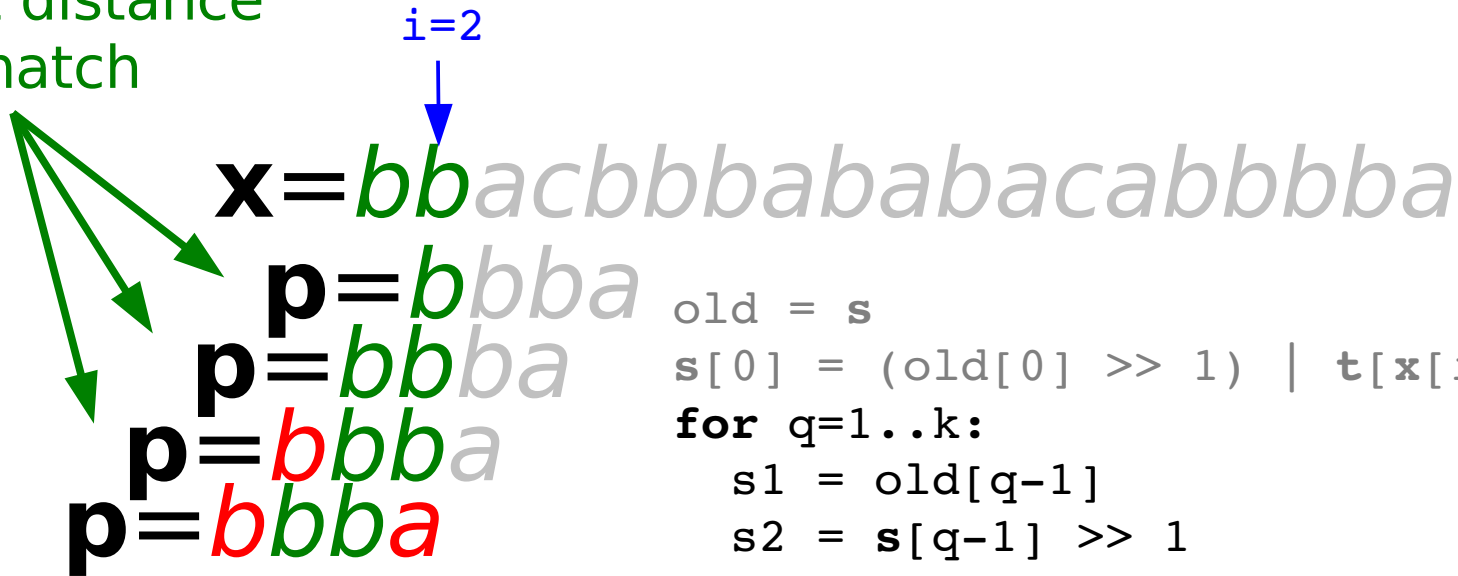
$i=2$   
↓  
 $x = bbacbbbababacabbba$   
 $p = bbba$   
 $p = bbba$   
 $p = bbba$   
 $p = bbba$

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
$s^1[0]:$	00111	$s^2[0]:$	00011	
$s^1[1]:$	00011	$s^2[1]:$	00001	= 00111 & 00001
$s^1[2]:$	00001	$s^2[2]:$		& 00011 & (00001   00001)

# Example

1 edit distance  
match



```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>		
$s^1[0]:$	00111	$s^2[0]:$	00011	=	00111 & 00001
$s^1[1]:$	00011	$s^2[1]:$	<u>0000</u> 1	&	00011 & (00001   00001)
$s^1[2]:$	00001	$s^2[2]:$			

# Example

Not 1 edit  
distance  
match

$i=2$

$x = bbacbbbababacabbba$

$p = bbba$

$p = bbba$

$p = bbba$

$p = bbba$

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
$s^1[0]:$	00111	$s^2[0]:$	00011	
$s^1[1]:$	00011	$s^2[1]:$	0000 <u>1</u>	= 00111 & 00001
$s^1[2]:$	00001	$s^2[2]:$		& 00011 & (00001   00001)

# Example

i=2  
↓

**x** = *bb*acbbbababacabbba

**p** = *bbba*

**p** = *bbba*

**p** = *bbba*

**p** = *bbba*

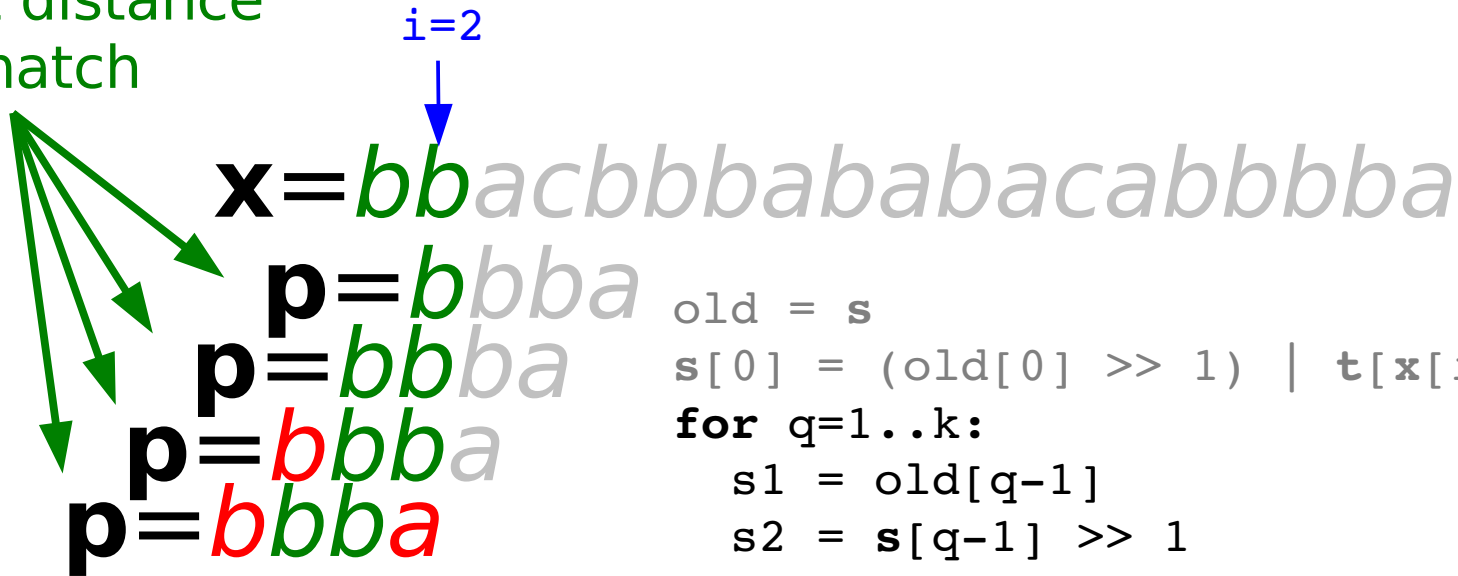
```

old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
  
```

	01234		01234	
$s^1[0]$ :	00111	$s^2[0]$ :	00011	
$s^1[1]$ :	00011	$s^2[1]$ :	00001	
$s^1[2]$ :	00001	$s^2[2]$ :	00000	= 00011 & 00000
				& 00001 & (00000   00001)

# Example

2 edit distance  
match

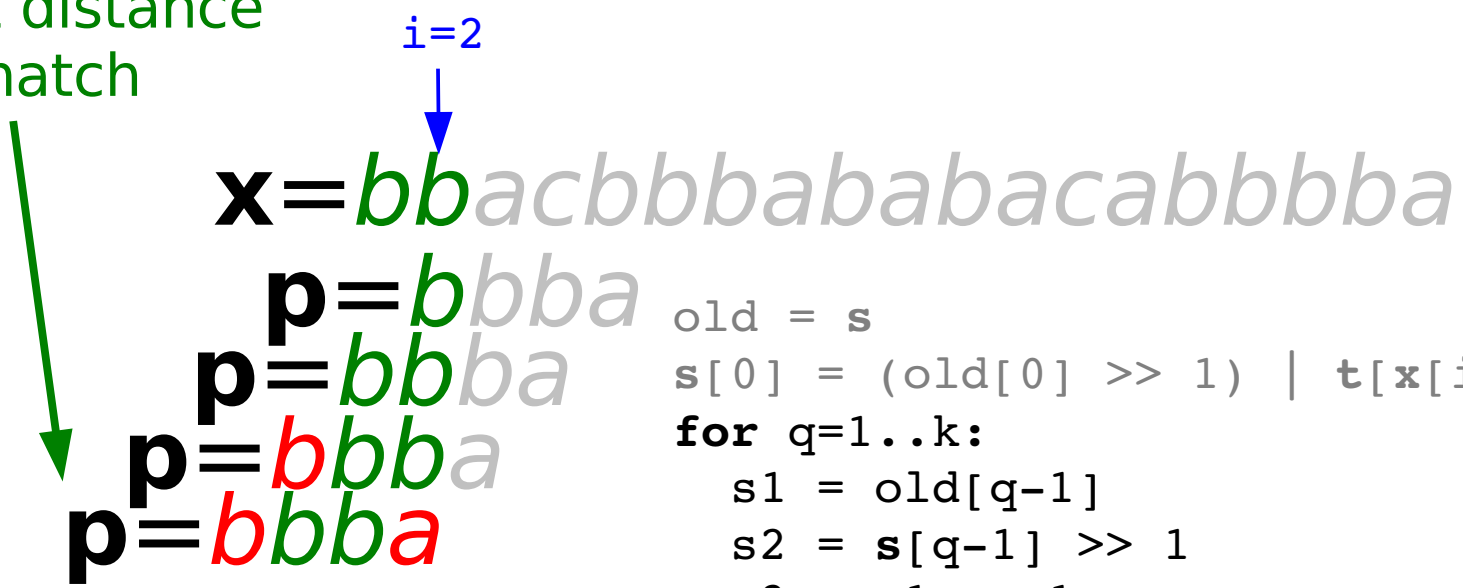


```
old = s  
s[0] = (old[0] >> 1) | t[x[i]]  
for q=1..k:  
  s1 = old[q-1]  
  s2 = s[q-1] >> 1  
  s3 = s1 >> 1  
  s4 = old[q] >> 1  
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

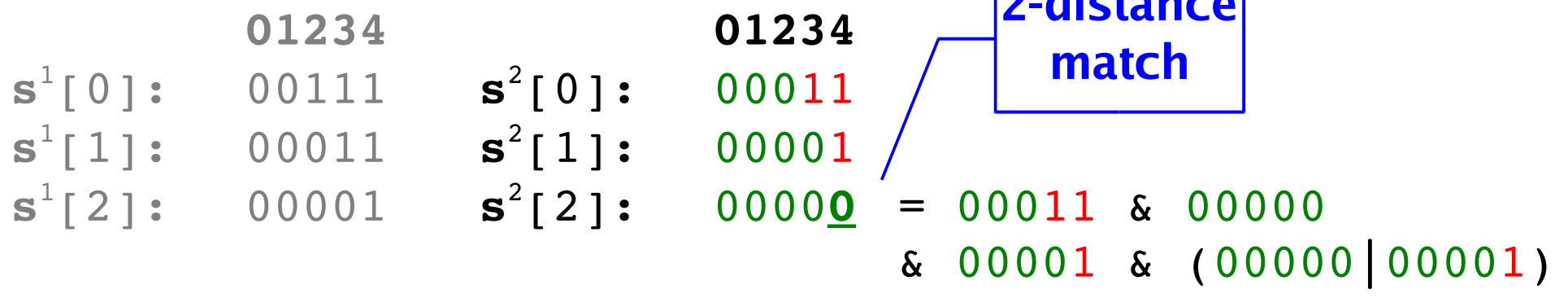
	<b>01234</b>		<b>01234</b>		
$s^1[0]:$	00111	$s^2[0]:$	00011	=	00011 & 00000
$s^1[1]:$	00011	$s^2[1]:$	00001	&	00001 & (00000   00001)
$s^1[2]:$	00001	$s^2[2]:$	<u>00000</u>		

# Example

2 edit distance match



```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```



# Example

i=3



**x**=*bb***a***cbbbbababacabbbba*

**p**=*b***b***bba*

**p**=*bb***b***ba*

**p**=*bbb***a**

**p**=*bbba*

```
old = s
```

```
s[0] = (old[0] >> 1) | t[x[i]]
```

```
for q=1..k:
```

```
  s1 = old[q-1]
```

```
  s2 = s[q-1] >> 1
```

```
  s3 = s1 >> 1
```

```
  s4 = old[q] >> 1
```

```
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	0	1	2	3	4
$s^2[0]:$	0	0	0	1	1
$s^2[1]:$	0	0	0	0	1
$s^2[2]:$	0	0	0	0	0

	0	1	2	3	4
$s^3[0]:$	0	1	1	1	1
$s^3[1]:$					
$s^3[2]:$					

$01111 = 00001 | 01110$

# Example

Not 0 edit  
distance  
match

i=3



**x** = *bb***a***cbbbbababacabbbba*

**p** = *bb***b***ba*

**p** = *bb***b***ba*

**p** = *bb***b***ba*

**p** = *bb***b***ba*

old = s

s[0] = (old[0] >> 1) | t[x[i]]

for q=1..k:

s1 = old[q-1]

s2 = s[q-1] >> 1

s3 = s1 >> 1

s4 = old[q] >> 1

s[q] = s1 & s2 & s3 & (s4 | t[x[i]])

**s**<sup>2</sup>[0]: 01234  
00011  
**s**<sup>2</sup>[1]: 00001  
**s**<sup>2</sup>[2]: 00000

**s**<sup>3</sup>[0]: 01234  
01111 = 00001 | 01110  
**s**<sup>3</sup>[1]:  
**s**<sup>3</sup>[2]:

# Example

i=3



**x**=*bb***a***cbbbbababacabbbba*

**p**=*bb***a***ba*  
**p**=*bb***a***ba*  
**p**=*bb***a***ba*  
**p**=*bb***a***ba*

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
<b>s</b> <sup>2</sup> [0]:	00011	<b>s</b> <sup>3</sup> [0]:	<b>01111</b>	
<b>s</b> <sup>2</sup> [1]:	00001	<b>s</b> <sup>3</sup> [1]:	<b>00000</b>	= <b>00001</b> & <b>00111</b>
<b>s</b> <sup>2</sup> [2]:	00000	<b>s</b> <sup>3</sup> [2]:		& <b>00000</b> & ( <b>00000</b>   <b>01110</b> )

# Example

1 edit distance match

$i=3$



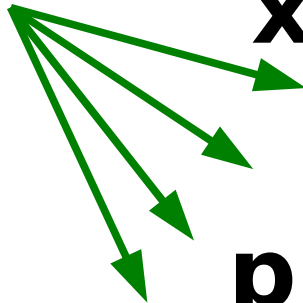
$x =$ *bb***a***cbbbbababacabbbba*

$p =$ *bb***b***ba*

$p =$ *b***b***ba*

$p =$ *bb***b***a*

$p =$ *b***bb***a*



`old = s`

`s[0] = (old[0] >> 1) | t[x[i]]`

`for q=1..k:`

`s1 = old[q-1]`

`s2 = s[q-1] >> 1`

`s3 = s1 >> 1`

`s4 = old[q] >> 1`

`s[q] = s1 & s2 & s3 & (s4 | t[x[i]])`

	<b>01234</b>		<b>01234</b>	
$s^2[0]:$	00011	$s^3[0]:$	<u>01111</u>	
$s^2[1]:$	00001	$s^3[1]:$	<u>00000</u>	$= 00001 \ \& \ 00111$
$s^2[2]:$	00000	$s^3[2]:$		$\& \ 00000 \ \& \ (00000 \   \ 01110)$

# Example

i=3



**x**=*bb***a***cbbbbababacabbbba*

**p**=*bb***a***cbbbbababacabbbba*  
**p**=*bb***a***cbbbbababacabbbba*  
**p**=*bb***a***cbbbbababacabbbba*  
**p**=*bb***a***cbbbbababacabbbba*

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
  s1 = old[q-1]
  s2 = s[q-1] >> 1
  s3 = s1 >> 1
  s4 = old[q] >> 1
  s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

	<b>01234</b>		<b>01234</b>	
<b>s<sup>2</sup>[0]:</b>	00011	<b>s<sup>3</sup>[0]:</b>	0 <b>1111</b>	
<b>s<sup>2</sup>[1]:</b>	00001	<b>s<sup>3</sup>[1]:</b>	00000	
<b>s<sup>2</sup>[2]:</b>	00000	<b>s<sup>3</sup>[2]:</b>	00000	= 0000 <b>1</b> & 00000
				& 00000 & (00000   0000 <b>1</b> )

# Example

2 edit distance  
match

i=3



**x** = *bb***a***cbbbbababacabbbba*

**p** = *b***bb*ba***

**p** = *bb***b***ba*

**p** = *bbb***a**

**p** = *bbba*

old = s

s[0] = (old[0] >> 1) | t[x[i]]

for q=1..k:

s1 = old[q-1]

s2 = s[q-1] >> 1

s3 = s1 >> 1

s4 = old[q] >> 1

s[q] = s1 & s2 & s3 & (s4 | t[x[i]])

01234  
s<sup>2</sup>[0]: 00011  
s<sup>2</sup>[1]: 00001  
s<sup>2</sup>[2]: 00000

01234  
s<sup>3</sup>[0]: 01111  
s<sup>3</sup>[1]: 00000  
s<sup>3</sup>[2]: 00000

= 00001 & 00000  
& 00000 & (00000 | 00001)

# Example

2 edit distance match

i=3

x = **bb**acbbbababacabbba

p = **b**bba  
 p = **bb**ba  
 p = **bbb**a  
 p = **bbba**

```
old = s
s[0] = (old[0] >> 1) | t[x[i]]
for q=1..k:
    s1 = old[q-1]
    s2 = s[q-1] >> 1
    s3 = s1 >> 1
    s4 = old[q] >> 1
    s[q] = s1 & s2 & s3 & (s4 | t[x[i]])
```

01234  
 s<sup>2</sup>[0]: 00011  
 s<sup>2</sup>[1]: 00001  
 s<sup>2</sup>[2]: 00000

01234  
 s<sup>3</sup>[0]: 0**1111**  
 s<sup>3</sup>[1]: 00000  
 s<sup>3</sup>[2]: 00000

2-distance match

$$= 00001 \& 00000$$

$$\& 00000 \& (00000 | 00001)$$

# Exercise:

---

---

Complete the example...

# Mandatory project 3

---

- In the final project you will
  - Use SHIFT-and-OR for exact matching
  - The SHIFT-and-OR variant for k-mismatch
  - Wu-Manber for edit-distance