

algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

scorrimento della stringa **testo** e matching
con la stringa **chiave**

A	Q	E	V	O	E	B	Q	O	E	A	Q	E	B	B	Q	O	E	A
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

B	Q	O	E
---	---	---	---

scorrere sempre di 1 ?

algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

scorrimento della stringa **testo** e matching
con la stringa **chiave**

A	Q	E	V	O	E	B	Q	O	E	A	Q	E	B	B	Q	O	E	A
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

B	Q	O	E
---	---	---	---

idea : considerare il primo carattere
del **testo** "a destra" della **chiave**

algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

scorrimento della stringa **testo** e matching
con la stringa **chiave**

A	Q	E	V	O	E	B	Q	O	E	A	Q	E	B	B	Q	O	E	A
---	---	---	---	----------	---	---	---	---	---	---	---	---	---	---	---	---	---	---

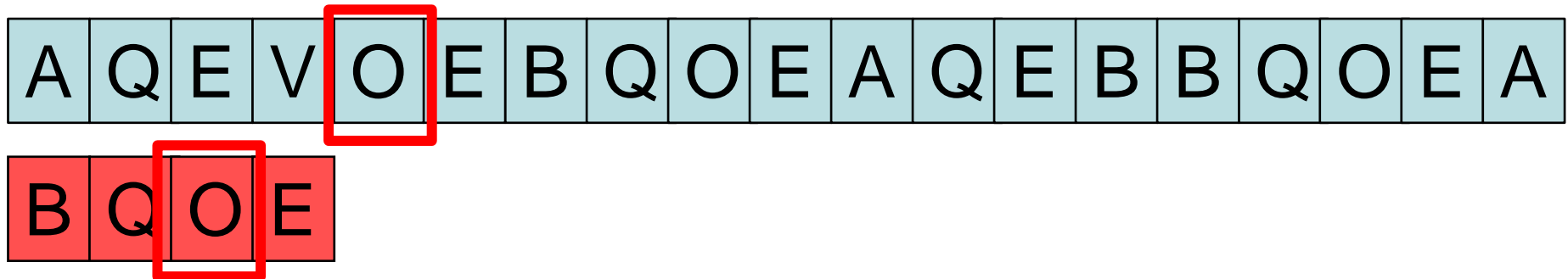
B	Q	O	E
----------	----------	----------	----------

idea : allineare opportunamente la **chiave**

algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

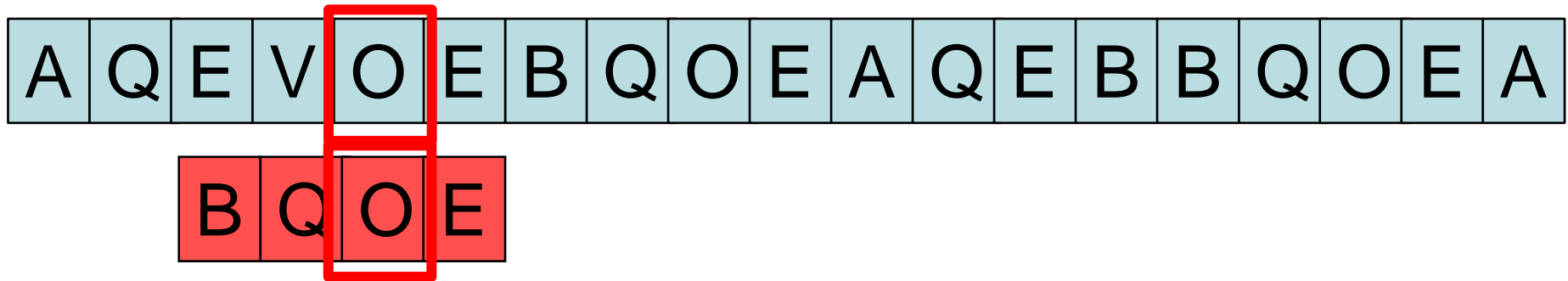
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

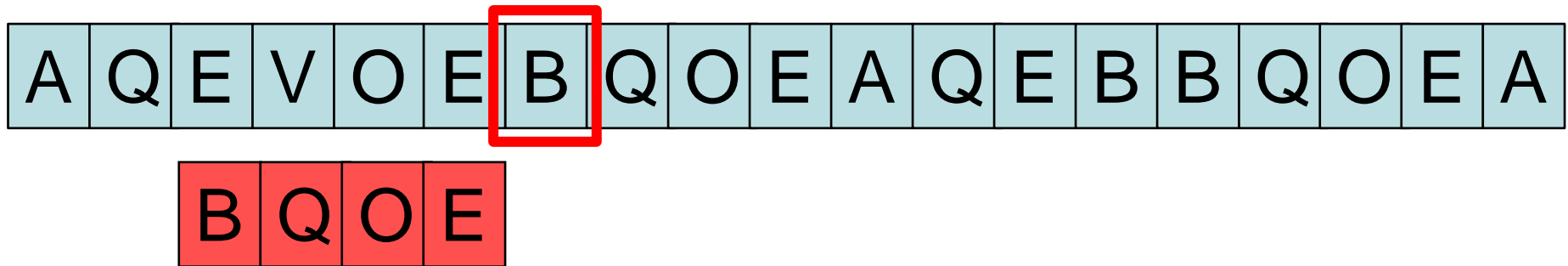
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

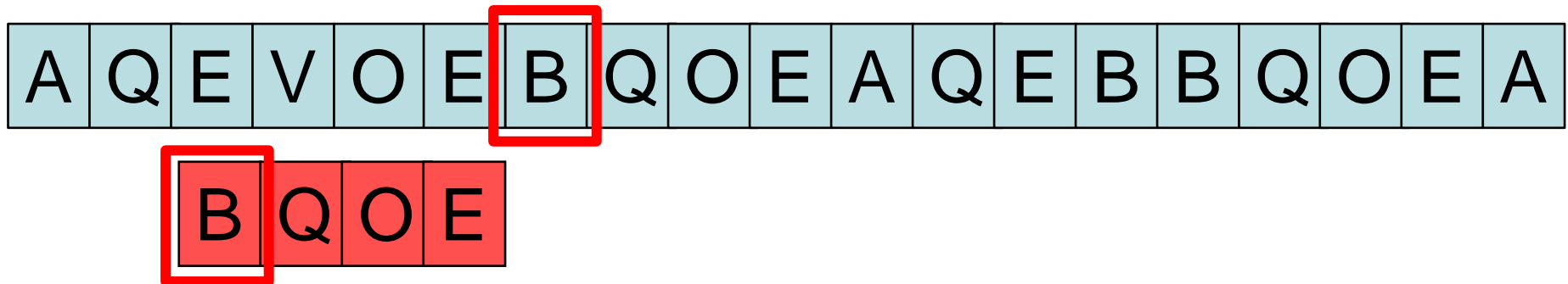
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

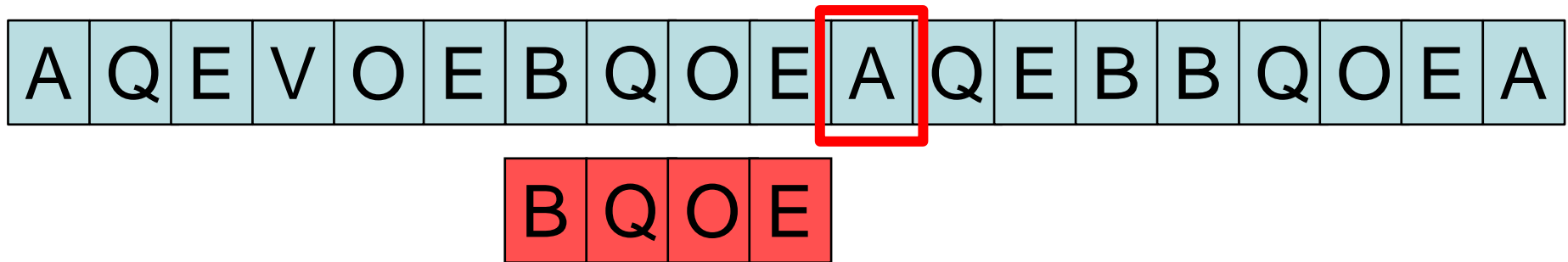
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

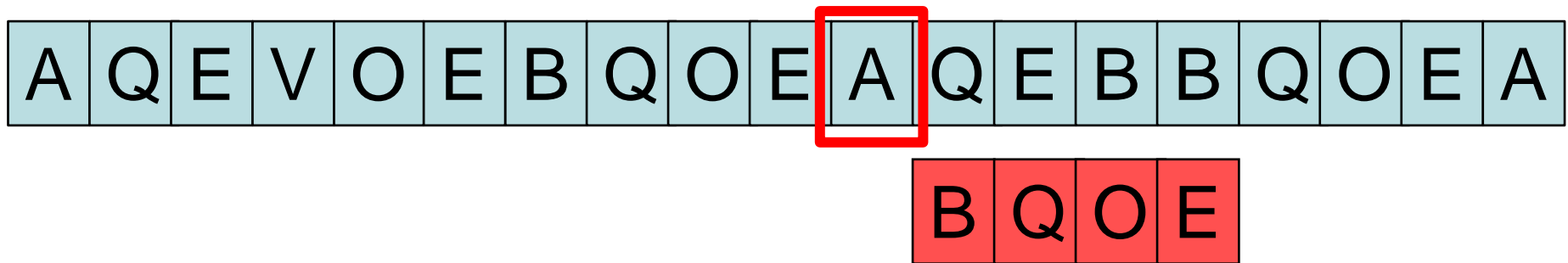
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

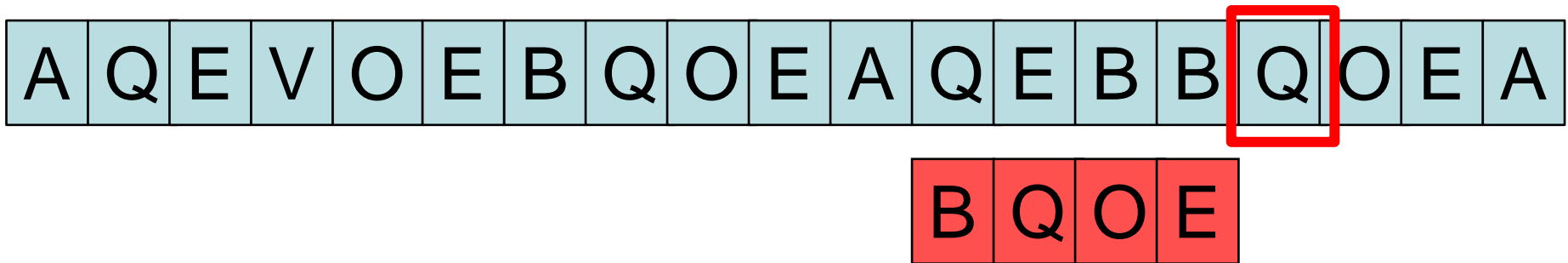
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

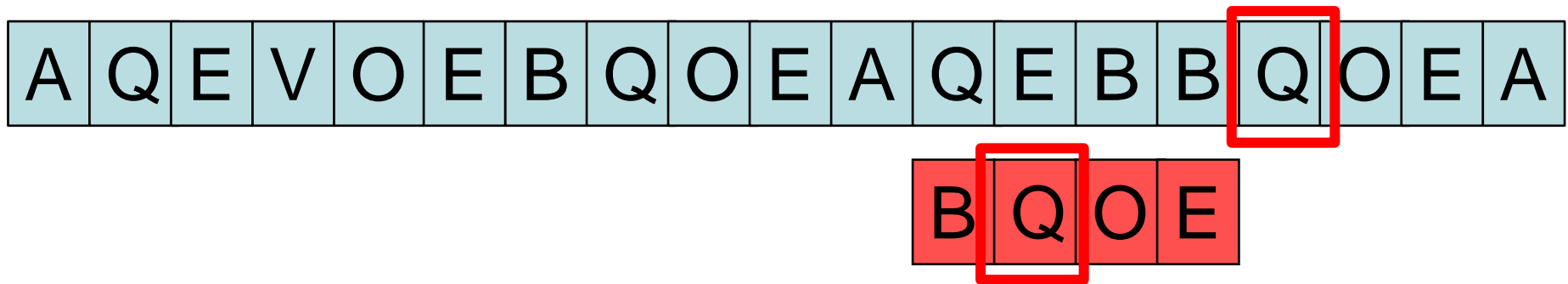
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

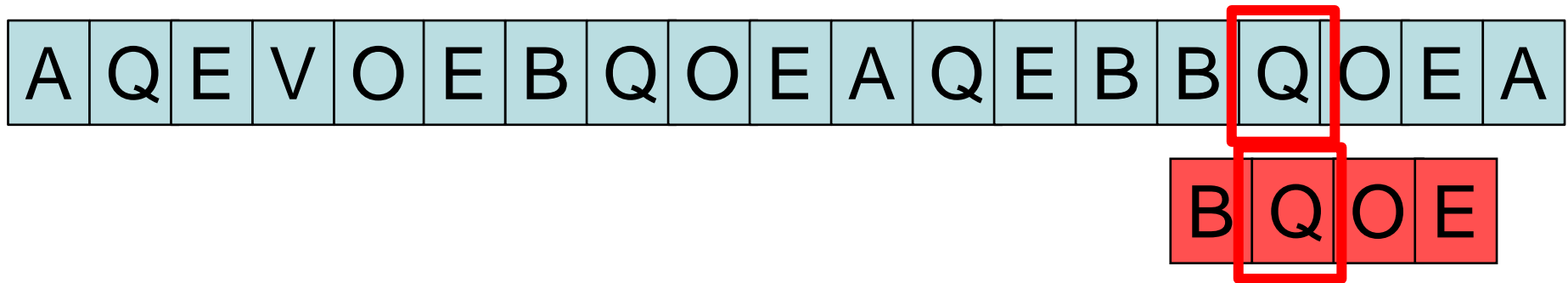
scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

scorrimento della stringa **testo** e matching
con la stringa **chiave**



algoritmo veloce di **ricerca** di
una stringa **chiave** in una stringa **testo**

versione più efficiente
dell'algoritmo di string-matching

- ✓ costruzione della **tabella degli shift** della stringa **chiave**
- ✓ **scorrimento** della stringa **testo** e matching con la stringa **chiave**

costruzione della **tabella degli shift** della stringa **chiave**

- determinare la lunghezza **m** della stringa **chiave**
- la tabella ha un numero di elementi pari al numero di caratteri dell'**alfabeto**
- sia **x** un carattere dell'alfabeto:
- se **x appartiene** alla stringa **chiave**
 - determinare la posizione **k** dell'**ultima** occorrenza di **x** nella stringa **chiave**
 - inserire nella tabella, nella posizione che corrisponde al carattere **x**, il numero **m-k**
- se **x non appartiene** alla stringa **chiave**
 - inserire nella tabella, nella posizione che corrisponde al carattere **x**, il numero **m+1**

A Q E V O E B Q O E A Q E B B Q O E A

B Q O E

$m = 4$

alfabeto

A, B, E, O, Q, V

tabella
degli shift

5, 4, 1, 2, 3, 5

se x **appartiene** alla stringa **chiave**

- k indice **ultima** occorrenza di x nella **chiave**
- in tabella $m-k$

se x **non appartiene** alla stringa **chiave**

- in tabella $m+1$

Esempio:

costruzione della **tabella degli shift** della **chiave**

alfabeto: A, C, G, T;

stringa **chiave** = CCGAGAG
0, 1, 2, 3, 4, 5, 6

lunghezza stringa chiave: 7 $m = 7$

tabella degli shift

	A	C	G	T
shift	2	6	1	8

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G	C	G	C	C	G	A	G	A	G	T	A	G	A	G	A	G	T	A	C	G
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

C	C	G	A	G	A	G
---	---	---	---	---	---	---

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

↕ mismatch

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

shift[G]=1, shift=1

G C G C C G A **G** A G T A G A G A G T A C G

↕ mismatch

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

↕ ↕ **mismatch**

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

↕ ↕ ↕ ↕ ↕ ↕ ↕ **exact match**

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

shift[T]=8, shift=8

G C G C C G A G A G **T** A G A G A G T A C G

↕ ↕ ↕ ↕ ↕ ↕ ↕ **exact match**

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T **A** G A G A G T A C G

↕ mismatch

C C G A G A G

uso della **tabella degli shift** nell'allineamento

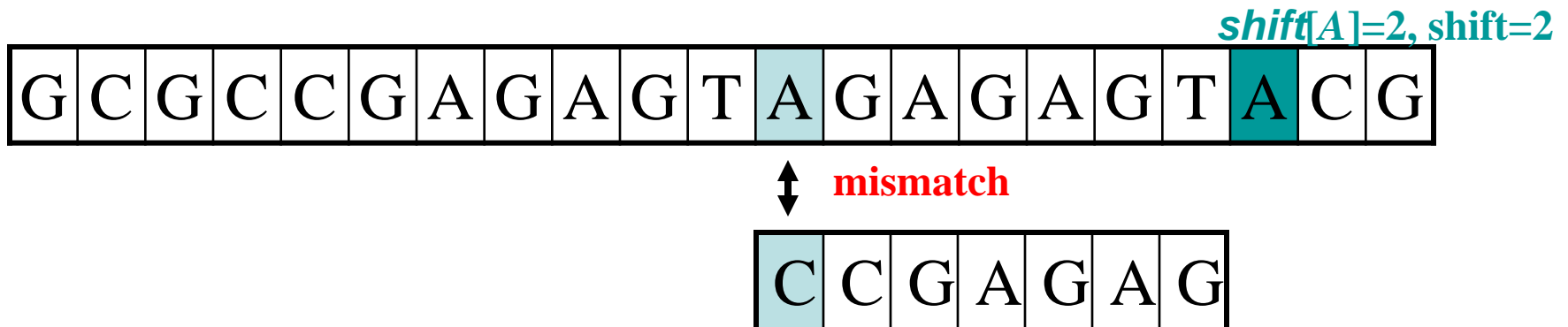
- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8



uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G	C	G	C	C	G	A	G	A	G	T	A	G	A	G	T	A	C	G
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

C	C	G	A	G	A	G
---	---	---	---	---	---	---

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C A G A G A G T A G **A** G A G T A C G

↕ mismatch

C A G A G A G

uso della **tabella degli shift** nell'allineamento

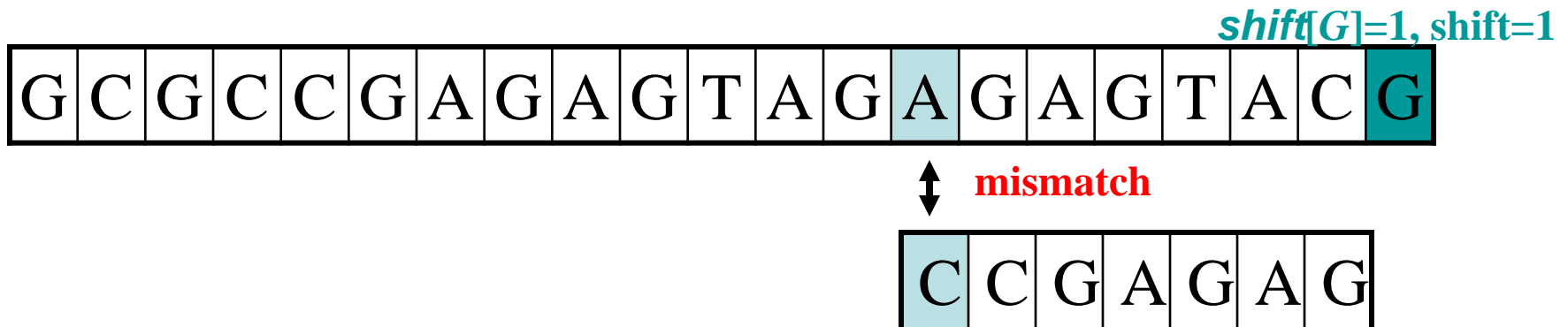
- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8



uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

C C G A G A G

uso della **tabella degli shift** nell'allineamento

- stringa testo

GCGCCGAGAGTAGAGAGTACG

- stringa chiave

CCGAGAG

	A	C	G	T
shift	2	6	1	8

G C G C C G A G A G T A G A G A G T A C G

↕ mismatch

C C G A G A G

Esercizio:

scrivere una function `tabshift` che riceve in **input**

- un `alfabeto`
- una stringa `chiave`

e restituisce in **output**

- la `tabella degli shift` della stringa `chiave`

scrivere una function `string_match_veloce` che riceve in **input**

- una stringa `testo`
- una stringa `chiave`

e restituisce in **output**

- il numero delle occorrenze di `chiave` in `testo`

Esercizio:

la function `string_match_veloce` deve

- ✓ costruire un array che contiene l'alfabeto,
- ✓ richiamare la function `tabshift` per costruire la tabella degli shift della chiave,
- ✓ implementare l'algoritmo di string matching veloce usando la tabella degli shift

usare la function `string_match_veloce` per determinare il numero delle occorrenze in stringa `testo`

`GCGCCGAGAGTAGAGAGTACG`

della stringa `chiave`

`CCGAGAG`