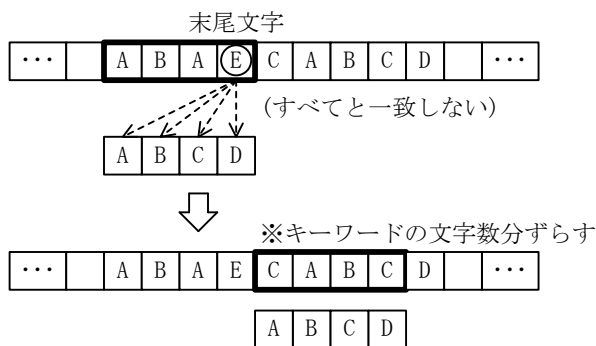


### 14-3. ボイヤ・ムーア法

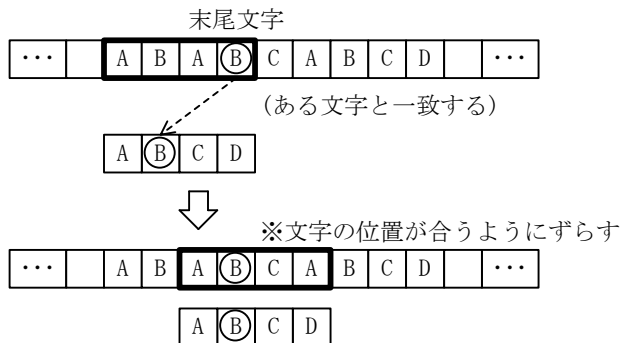
**ボイヤ・ムーア法**は、このアルゴリズムを開発したBoyer（ボイヤー）とMoore（ムーア）の名前をつなげた探索アルゴリズムである。ボイヤ・ムーア法は、対象文字列から抽出した部分文字列の末尾文字がキーワード中に存在しているかに着目し、行う必要のない無駄な比較を省略する方法である。

#### 【ボイヤ・ムーア法の考え方】

- (1) 部分文字列の末尾文字がキーワード中に存在しない場合、その末尾文字を含む部分文字列とキーワードを比較しても、絶対に一致しない。したがって、その末尾文字以降が次の部分文字列となるように、部分文字列をキーワードの文字数分ずらす。



- (2) 部分文字列の末尾文字がキーワード中に存在する場合、その二つの文字の位置を合わせなければ部分文字列とキーワードは、絶対に一致しない。したがって、二つの文字が同じ位置になるように部分文字列の位置をずらす。



ボイヤ・ムーア法では、部分文字列の末尾文字によって、部分文字列を移動する距離（移動文字数）が決定する。移動文字数はキーワード中に存在する文字によって異なるため、部分文字列の末尾文字がキーワード中の文字 $K[J]$ と一致した場合の移動文字数を $M[J]$ と表せるように、あらかじめ配列 $M$ に移動文字数を求めておく必要がある。

移動文字数の基本的な求め方を、配列Kに格納されたキーワード（Klen文字）を利用して説明する。

配列K 

A	B	C	D			
---	---	---	---	--	--	--

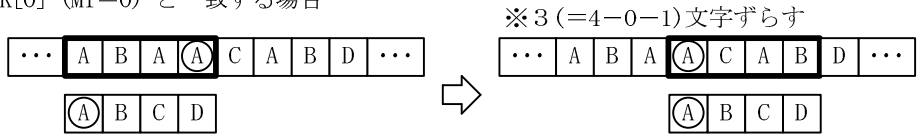
 ...キーワードが格納された配列  
 Klen 

4
---

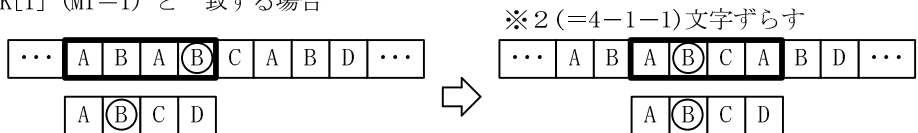
 ...キーワードの文字数

1) 部分文字列の末尾文字がK[M1] (M1=0~Klen-2) と一致する場合、移動文字数は (Klen-M1-1) 文字となる。

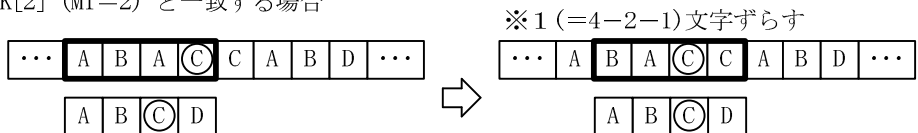
・K[0] (M1=0) と一致する場合



・K[1] (M1=1) と一致する場合

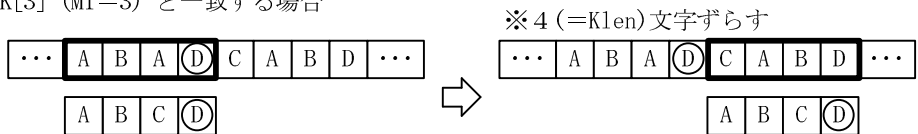


・K[2] (M1=2) と一致する場合



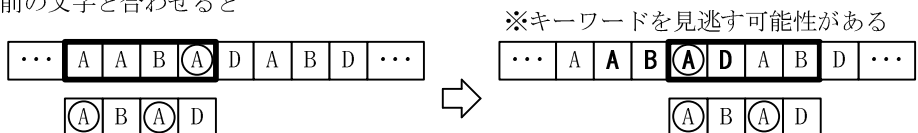
2) 部分文字列の末尾文字がキーワードの末尾文字 (K[Klen-1]) と一致する場合、移動文字数はキーワードの文字数分 (Klen文字) となる。

・K[3] (M1=3) と一致する場合



3) キーワード中に同じ文字が複数ある場合、末尾に最も近い文字と合わせるように移動文字数を決定する。

・前の文字と合わせると



・後の文字と合わせると

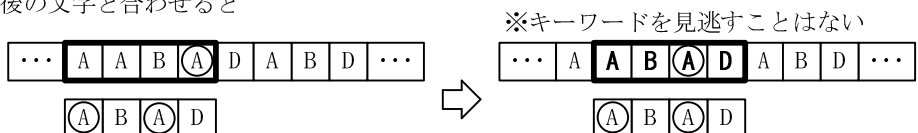


図14-9は、部分文字列の末尾の文字がキーワードのK[J]と一致したときの移動文字数を、配列MのM[J]に求める流れ図である。なお、この流れ図は、あとでボイヤ・ムーア法の流れ図から呼び出すために、定義済み処理として記述してある。

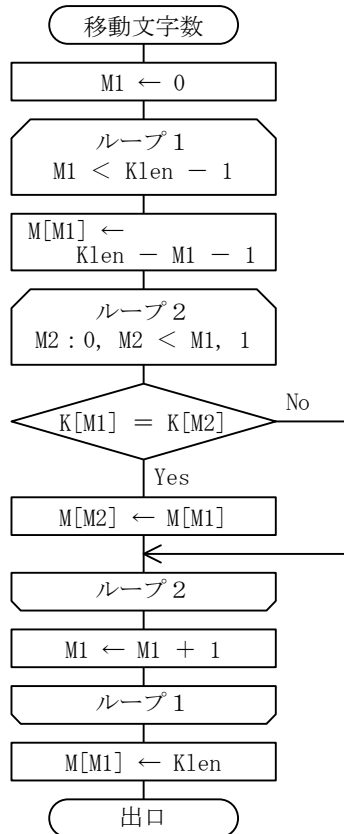
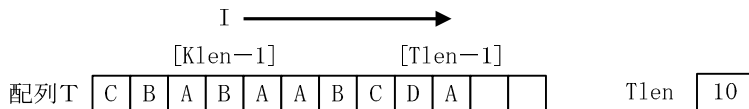


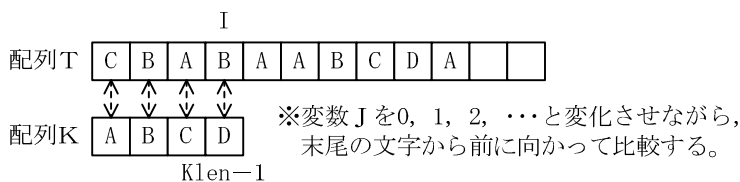
図14-9 移動文字数を求める流れ図

この移動文字数が求められると、ボイヤ・ムーア法は比較的簡単に実現できる。

まず、配列Tに格納された対象文字列（Tlen文字）の比較対象文字を、変数Iを用いてT[I]で表す。このとき、部分文字列の末尾文字から比較を行うので、変数Iの初期値は（Klen-1）となり、対象文字列の最後の文字を表す（Tlen-1）まで変化する。



次に、部分文字列とキーワードの文字比較は、変数Iの位置を動かしたくないので、変数Jを0～（Klen-1）と変化させ、T[I-J]とK[Klen-J-1]の比較で実現する。



最後に、探索成功時に出力するキーワードと一致した部分文字列の開始位置（添字）は、変数  $I$  が部分文字列の末尾文字の添字なので  $(I - Klen + 1)$  で求める。

以上の考え方をまとめたものが、図14-10の流れ図である。このアルゴリズムの効率の良さを確認するために、データを変えながら実際にトレースをしてもらいたい。

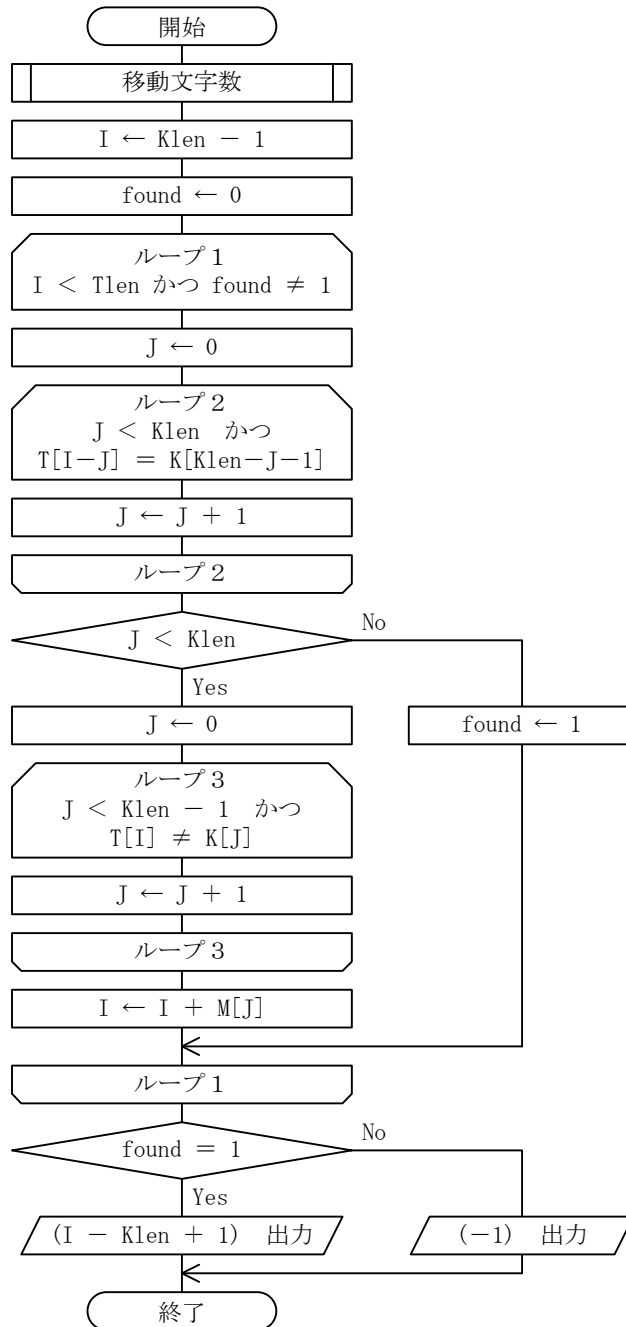


図14-10 ボイヤ・ムーア法の流れ図

# 問14-4

配列 T に格納されている N 文字の探索文字列から、配列 K に格納されている M 文字のキーワードをボイヤ・ムーア法で探索し、一致した部分文字列の先頭の添字を出力する流れ図中の空欄 a ~ d に入れる適切な処理を答えよ。なお、キーワードから求めた移動文字数は、要素数 M 個の配列 S にあらかじめ求められているものとし、キーワードと一致する部分文字列が見つからない場合は (-1) を出力するものとする。

配列 T 

A	B	A	B	D	A	B	A	B	A	C	D	A	B					
---	---	---	---	---	---	---	---	---	---	---	---	---	---	--	--	--	--	--

 N 

14
----

配列 K 

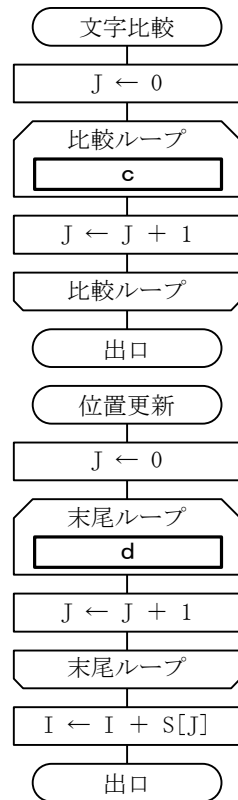
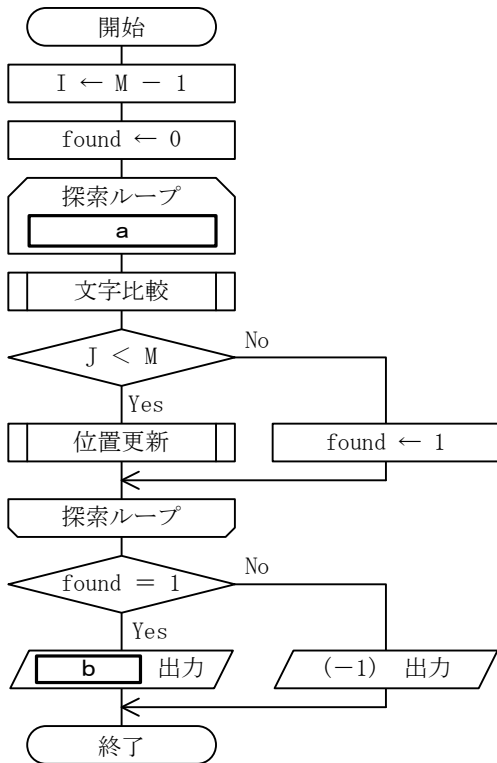
A	B	A	C	D
---	---	---	---	---

 配列 S 

2	3	2	1	5
---	---	---	---	---

 M 

5
---



a	
b	
c	
d	