

Problem 3 :

子題 1：出現的次數最多的 DNA 序列

給予一堆相同長度的 ATCG 構成的基因，對測試資料的字串進行分析，對於每一個位置，找出 ATCG 字符使用最多次的基因，得到一個最優的基因字串，使它滿足與測試資料的基因字串不同的字元之和最小。如果有多種解，則輸出字典序最小(A,C,G,T)的基因。

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

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

第一行 ACGT 和最後一行 CGTA，ATCG 構成的基因，4 個都出現 1 次，輸出字典序最小(A,C,G,T)的基因，所以這二行輸出 A。

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

最後一行 TTATAA，ATCG 構成的基因，TA2 個都出現 3 次，輸出字典序最小(A,C,G,T)的基因，所以這一行輸出 A。

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

第一行 GCGC、第二行 AACC、第三行 GTTG 和第四行 CCTT，ATCG 構成的基因，各出現 2 次，輸出字典序最小(A,C,G,T)的基因，所以這四行輸出分別為 CAGC。最後一行 CGTA，ATCG 構成的基因，4 個都出現 1 次，輸出字典序最小(A,C,G,T)的基因，所以這一行輸出

A。

輸入說明：

第一列的數字 n 代表有幾筆資料要測試， $2 \leq n \leq 20$ ，之後為每組的測試資料，每組的測試資料，第一列的二個數字 $A B$ ， $4 \leq A, B \leq 50$ ，以一個或多個空白作為區隔。之後有 A 個長度為 B 的 DNA 序列，直接統計每位四個字母出現的次數，找到出現最多的就是解。如果個數相同取字典序小(A,C,G,T)的基因。

輸出說明：

每筆測試資料輸出一列。出現的次數最多的 DNA 序列，長度為 B 的 DNA 序列。

輸入檔案 1：【檔名：in1.txt】

```
2
5 8
TATGATAC
TAAGCTAC
AAAGATCC
TGAGATAC
TGAGATAC
4 10
ACGTACGTAC
CCGTACGTAG
GCGTACGTAT
TCGTACGTAA
```

輸入檔案 2：【檔名：in2.txt】

```
2
6 10
ATGTTACCAT
AAGTTACGAT
AACAAAGCAA
AAGTTACCTT
AAGTTACCAA
TACTTACCAA
4 9
GAGCACGTC
CATCACGTG
GCTTACGTT
CCGTATGTA
```

輸出範例：【檔名：out1.txt】

TAAGATAC

ACGTACGTAA

輸出範例：【檔名：out2.txt】

AAGTTACCAA

CAGCACGTA

資料來源：Uva1368