

문제	1-A
제목	DNA 대표서열
내용	<div data-bbox="630 425 1061 1041" data-label="Chemical-Block"> </div> <p>그림과 같이 DNA는 유전정보를 저장하고 있는 물질이다. DNA는 아데닌(Adenine), 티민(Thymine), 구아닌(Guanine), 시토신(Cytosine) 등 4가지의 염기(nucleotide)로 구성되어 있다. DNA를 구성하는 각 염기들의 첫 문자들로 DNA를 표현하면 A, C, G, T 4개의 문자로 구성된 문자열이 된다. 생물학 교수 김인하는 동일한 유전자(gene)가 여러 종의 동물에 존재하며 종마다 동일유전자의 DNA 서열이 조금씩 다를 수 있다는 것을 알게 되었다.</p> <p>예를 들어 고양이, 개, 말, 소, 원숭이에 공통적으로 존재하는 유전자 X의 서열이 아래와 같다. 이때 X의 DNA 서열은 각 종마다 동일하지는 않을 수 있지만 대단히 유사하다.</p> <p>유전자 X의 DNA 서열 고양이: GCATATGGCTGTGCA 개: GCAAATGGCTGTGCA 말: GCTAATGGGTGTCCA 소: GCAAATGGCTGTGCA 원숭이: GCAAATCGGTGAGCA</p> <p>김인하 교수는 유전자 X가 인간에게도 존재할 수 있다고 생각하고 X를 검색하기 위해 5종의 동물에 존재하는 X의 서열을 대표할 수 있는 서열을 만들기 시작했다. 즉, 각 종의 유전자 서열과의 해밍거리의 합이 최소인 서열들 중, 알파벳 순서가 가장 빠른 서열을 대표서열로 정의하기로 했다. 여기에서 해밍거리는 길이가 같은 두 문자열의 동일한 위치에서 서로 다른 문자들의 수</p>

		<p>를 뜻한다. 예를 들어, 두 문자열 AGCAT와 GGAAT가 주어졌을 때, 첫 번째 위치의 문자가 각각 A와 G로 다르고, 세 번째 위치의 문자가 각각 C와 A로 다르므로 AGCAT와 GGAAT의 해밍거리는 2이다. 김인하 교수를 돕기 위해, 유전자 X에 대한 여러 종의 DNA 서열이 주어졌을 때, 유전자 X의 대표서열을 출력하고, 대표서열과 각 종의 유전자 서열과의 해밍거리의 합을 출력하는 프로그램을 작성하시오.</p> <p>사용 할 수 있는 언어는 C, C++로 제한한다. 프로그램의 실행시간은 1초를 초과할 수 없다. C++의 경우 main 함수 내의 시작 지점에 다음 내용을 추가함으로써 cin 입력 속도를 개선할 수 있다.</p> <pre>std::ios::sync_with_stdio(false);</pre>
입력 형식		<p>첫 줄에는 테스트케이스의 수 $T(≤ 1000)$가 주어진다. 각 테스트케이스의 첫 번째 줄에는 종의 수 $M(1 ≤ M ≤ 50)$과 각 유전자의 길이 $N(1 ≤ N ≤ 1000)$이 공백 하나를 사이에 두고 주어진다. 각 테스트케이스의 두 번째 줄부터 M개의 줄에는 종별 유전자 X의 DNA 서열이 한 줄에 하나씩 주어진다. 이후에 같은 형태의 테스트케이스가 $T-1$개 더 반복된다.</p>
출력 형식		<p>출력은 표준출력으로 수행하며 총 $2T$줄로 이루어진다. 각 테스트케이스 별로 첫 번째 줄에는 대표서열을 출력하고, 두 번째 줄에는 대표서열과 각 종의 유전자 서열과의 해밍거리의 합을 출력한다.</p>
예	입력	<pre>3 5 8 TATGATAC TAAGCTAC AAAGATCC TGAGATAC TAAGATGT 4 10 ACGTACGTAC CCGTACGTAG GCGTACGTAT TCGTACGTAA 6 10 ATGTTACCAT AAGTTACGAT AACAAAGCAA AAGTTACCTT AAGTTACCAA TACTTACCAA (empty line)</pre>
	출력	<pre>TAAGATAC 7 ACGTACGTAA 6 AAGTTACCAA 12 (empty line)</pre>