

СЕДЕМНАДЕСЕТИ КОНКУРС „МИНКО БАЛКАНСКИ”
Състезание по информатика, София, 15 май 2010 г.

Задача 2. Компресия

Задачата Ви е да напишете програма **compression**, която компресира и декомпресира голямо количество текст. По-точно, трябва да създадете колкото е възможно по-къса битова последователност, от която след това да възстановите първоначалния текст. Всеки от предоставяните текстове притежава няколко специални свойства:

- всички символи са случайно разбъркани малки латински букви;
- дължината му L ($10^3 \leq L \leq 10^6$) е кратна на 100;
- някои букви могат да се срещат много по-често от други, но никоя буква не се среща повече от $L/2$ пъти.

Вход и изход:

На първия ред на стандартния вход е записана команда, указваща операцията, която трябва да извърши Вашата програма.

При команда “compress” трябва да прочетете файла “compression.in”, и да създадете файл “compression.cmp”, съдържащ единствено символи ‘0’ и ‘1’.

При команда “decompress” трябва да прочетете файла “compression.cmp” (който е създаден при предишно изпълнение на програмата Ви) и да създадете файл “compression.out”.

Всички файлове се интерпретират като текстови.

Оценяване:

Ако съдържанието на файловете “compression.in” и “compression.out” се различава, след като последователно на програмата Ви е било зададено да компресира и декомпресира файл, ще получите 0 точки за съответния тест. Ако работата на програмата Ви е коректна, за съответния тест ще получите $10 * [\min(\text{best}/\text{your}, 1)]^3$ точки, където *your* е броят на символите във файла, компресиран от Вашата програма, а *best* е минималният брой на символите във файл, компресиран от програмата на участник или автора.

Пример (с нарушение на правилото за минимална дължина):

compression.in	adccaababba
compression.cmp	1000000100000110000110000101001 0000001101101011
compression.out	adccaababba