

Kompresja

WWI 2024 – Off

Kod zadania: **zip**
Limit pamięci: **512 MiB**



Zadanie polega na zaprojektowaniu dobrego algorytmu kompresji tekstu w języku polskim. Twoim zadaniem jest napisanie dwóch funkcji:

- `std::string zip(const std::string &text)`, która przyjmuje tekst do skompresowania i zwraca skompresowany tekst.
- `std::string unzip(const std::string &compressed)`, która przyjmuje skompresowany tekst zwrócony przez `zip` i dekompresuje go.

Wejście i wyjście

Zarówno nieskompresowany, jak i skompresowany tekst mogą składać się wyłącznie z małych liter alfabetu łacińskiego oraz znaków podkreślenia (`_`) reprezentujących spacje. Nieskompresowany tekst będzie zawierał wyłącznie pojedyncze podkreślenia, ale w skompresowanym tekście podkreślenia mogą występować po kilka razy.

Każdy fragment tekstu będzie zawierał więcej niż 800 znaków i mniej niż 2^{16} .

Testy zostały wygenerowane mechanicznie z tekstów z różnych źródeł, więc oprócz słów w języku polskim mogą zawierać również śladowe ilości nazw własnych, słów w innych językach, literówek, skrótów, liczb rzymskich itp.

Wymagania techniczne

Twój program nie powinien zawierać funkcji `main`, ale może zawierać funkcje pomocnicze.

Zadanie **nie polega** na oszukaniu sprawdzaczki, więc wszelkie próby ataków side-channel itp. będą karane wyzerowaniem zgłoszenia lub dyskwalifikacją. Kompresujcie uczciwie!

Ocenianie

Niech x oznacza rozmiar danych przed kompresją, a y rozmiar danych po kompresji. Program otrzyma 0 punktów za dany test jeśli $y \geq x$. W przeciwnym razie otrzyma

$$100 \left(1 - \frac{y}{x} \right)$$

procent punktów przewidzianych za test.