

Opis problemu algorytmicznego

W kryptoanalizie, czyli nauce związanej z odgadywaniem szyfrów, w metodach szyfrów podstawieniowych i przestawieniowych, korzysta się z informacji, która pomaga złamać szyfr na zasadzie odkrycia liter z tekstu zaszyfrowanego na podstawie częstości ich występowania. Każdy język ma charakterystyczną dla siebie częstość występowania poszczególnych liter w tekście.

W języku angielskim statystycznie przyjmuje się następującą częstość występowania liter:

A	B	C	D	E	F	G	H	I	J	K	L	M
0,082	0,015	0,028	0,043	0,127	0,022	0,020	0,061	0,070	0,002	0,008	0,040	0,024
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0,067	0,075	0,019	0,001	0,060	0,063	0,091	0,028	0,010	0,023	0,001	0,020	0,001

W języku polskim statystycznie przyjmuje się następującą częstość występowania liter:

a 8.91%	w 4.65%	p 3.13%	g 1.42%	ć 0.40%
i 8.21%	s 4.32%	m 2.80%	ę 1.11%	f 0.30%
o 7.75%	t 3.98%	u 2.50%	h 1.08%	ń 0.20%
e 7.66%	c 3.96%	j 2.28%	ą 0.99%	q 0.14%
z 5.64%	y 3.76%	l 2.10%	ó 0.85%	ż 0.06%
n 5.52%	k 3.51%	ł 1.82%	ź 0.83%	v 0.04%
r 4.69%	d 3.25%	b 1.47%	ś 0.66%	x 0.02%

— Adam Przepiórkowski, Instytut Podstaw Informatyki PAN

Źródło: <https://sjp.pwn.pl/poradnia/haslo/frekwencja-liter-w-polskich-tekstach;7072.html>

Implementacja

Biorąc pod uwagę powyższe informacje należy napisać program, który dokonuje analizy tekstów o długości do 100 znaków i określa w jakim języku ten tekst jest napisany, biorąc pod uwagę tylko te dwa języki: polski i angielski.

Tekst wejściowy ma być przez użytkownika podawany z klawiatury.

Na ekranie ma być wyświetlana informacja o rodzaju języka oraz częstość poszczególnych liter w podanym przez użytkownika tekście w nawiązaniu do częstości podanej w materiale do zadania.

Zgodnie z dobrą praktyką programistyczną należy dodać w programie komentarze, objaśniające co wykonują poszczególne partie kodu.

Dodatkowo do rozwiązania w postaci programu (zgodnie z regulaminem), należy dołączyć dokument PDF, w którym uczestnik/zespół przedstawi/ą, jakie założenia algorytmiczne przyjął/przyjęli do napisania swojego programu.