

Лабораторне заняття 3.

Застосування алгоритмів ефективного кодування

Мета: отримати навички використання алгоритмів ефективного кодування

Завдання 1. Виконати ефективне кодування літер української абетки за допомогою алгоритму Шеннона-Фано

Відносні частоти (імовірності) вживання літер української абетки.

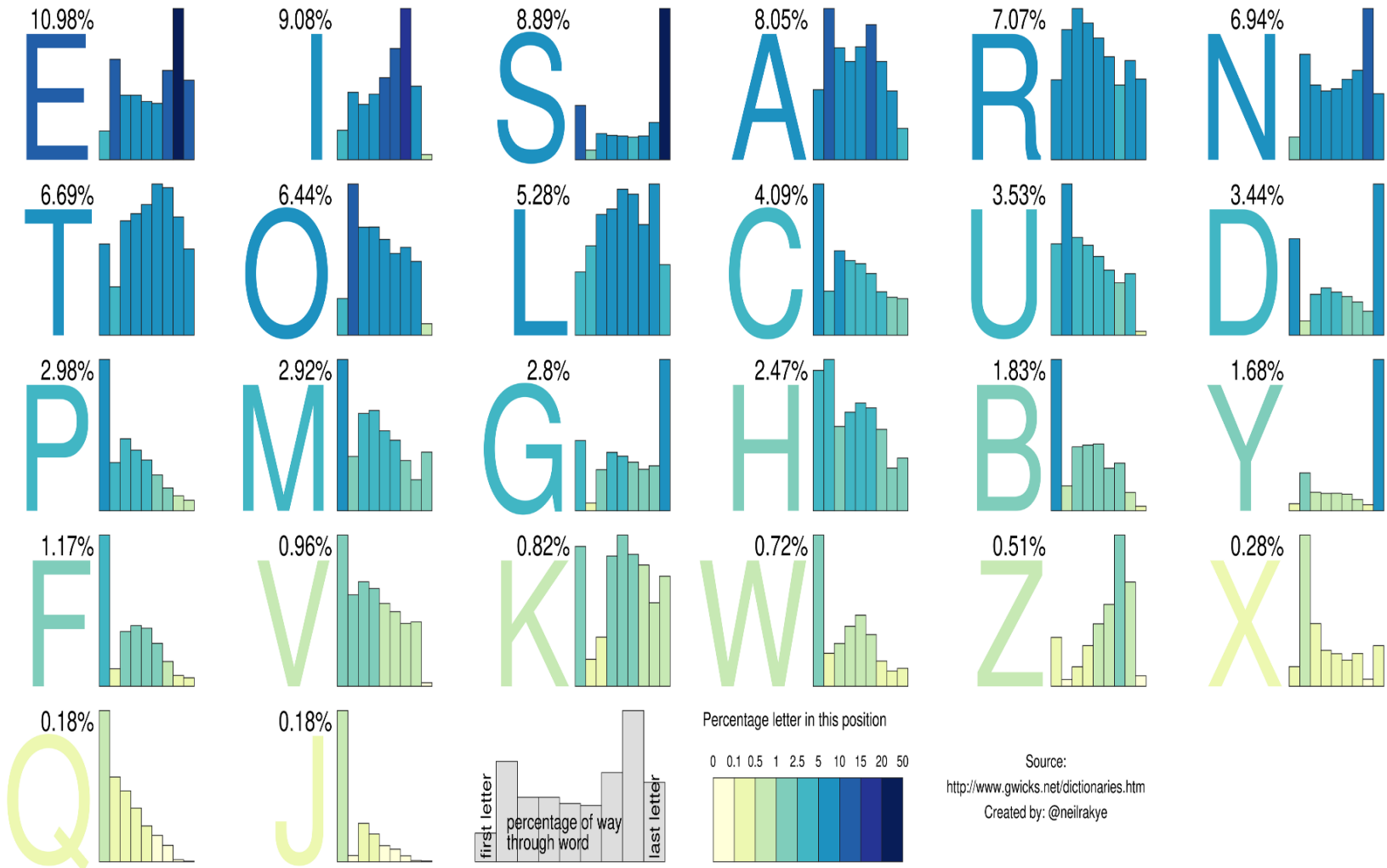
Літера	Частота	Літера	Частота	Літера	Частота	Літера	Частота
„—”	0,138	Р	0,043	П	0,025	Ї	0,010
О	0,086	Е	0,042	З	0,020	Ц	0,010
Н	0,068	С	0,037	Я	0,019	Й	0,009
А	0,064	К	0,033	Ь	0,016	Ю	0,008
И	0,055	М	0,029	Б	0,013	Ж	0,007
В	0,046	У	0,027	Г	0,013	Ш	0,005
Т	0,045	Д	0,027	Ч	0,011	Є	0,005
І	0,044	Л	0,027	Х	0,011	Щ	0,004
						Ф	0,003

Частота повторюваності букви Ґ дорівнює 0,00006

Рекомендація: доцільно застосувати електронну таблицю

Завдання 2. Виконати ефективне кодування літер англійського алфавіту за допомогою алгоритму Шеннона-Фано

Frequency of letters in English words and where they occur in the word



Порівняти отримані коди.

Індивідуальне завдання на самостійну роботу: виконати кодування літер абеток алгоритмом Хаффмана і порівняти результати з алгоритмом Шеннона-Фано.

Приклад

Який алгоритм використаний?

№ розряду Знаки	1	2	3	4	5	6	7	8	9	Код літери
–	0	0	0							000
О			1							001
Н		1	0	0						0100
А			1	0						0101
И			1	0						0110
В			1	0						0111
Т			0	0						1000
І	0	0	1	1001						
Р		1	0	0	10100					
Е			1	0	10101					
С			1	0	10110					
К			1	0	10111					
М		0	0	0	11000					
У			1	0	110010					
Д			0	0	0	110011				
Л				1	0	110100				
П			1	0	110110					
З	1		0	110111						
Я	1	0	0	0	111000					
Ь			1	0	111001					
Б			1	0	111010					
Г			1	0	111011					
Ч		1	0	0	111100					
Х			1	0	1111010					
І	1		0	1111011						
Ц	1	1	0	0	0	1111100				
И				1	0	1111111				
Ю				1	0	0	11111100			
Ж					1	0	11111101			
Ш	1	1	1	0	0	11111110				
Є				1	0	11111111				
Щ				0	0	111111100				
Ф				1	0	111111101				

Частота повторюваності букви Г в текстах українською мовою дорівнює $6 \cdot 10^{-5}$.