Комп'ютерний практикум №1 з Криптографії

Грипас Владислав ФБ-12

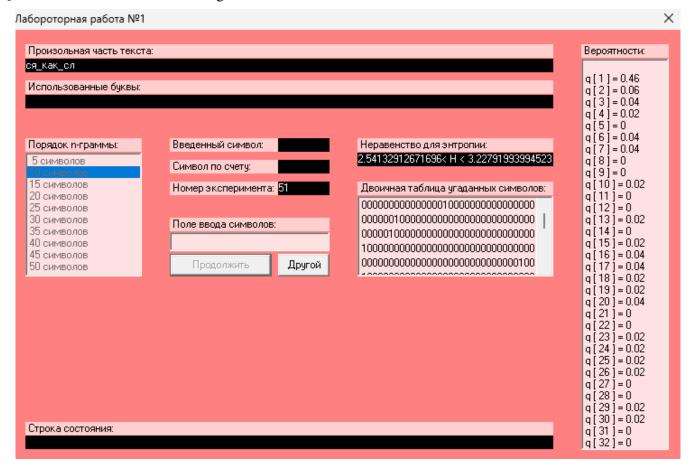
Мета роботи: Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення на порівняння рзіних моделей джерела відкритого тескту для наближеного визнчення ентропії, набуття практичних начок щодо оцінки ентропіх на символ джерела Постановка задачі:

- 1) Написати програму для підрахунку частот букв і біграм в тексті, а також підрахунку Н1 та Н2 за безпосереднім означенням. Підрахувати частити букв та біграм, а також занчення Н1 та Н2 надовільно обраному тексті російською мовою достатньої довжини, де імовірності замінити відповідними частотами. Також одержати значення Н1 та Н2 на тому ж тексі, де вилучені всі пробіли
- 2) За допомогою програми CoolPinkProgram оцінити значення H10, H20, H30
- 3) Використовуючи отримані значення ентропії, оцінити надлишковість російськох мови в різних моделях джерела

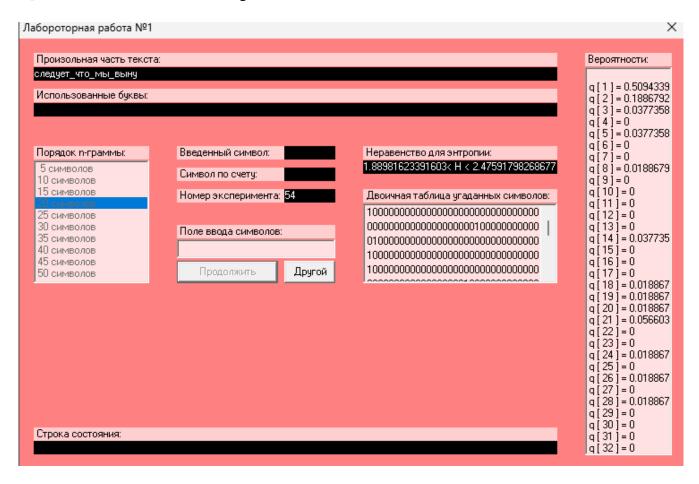
Хід роботи:

Мною був вибраний текстовий файл зі змістом книги "Брати Карамазови" Для початку я ініціалізую алфавіт, що буде використовуватись в подальших обрахунках, та створюю його копію, додаючи пробіл, далі, обрахувавши необхідні показники, виводжу головні з них в вивід терміналу, записуючи інші дані в файли з відповідними назвами При виконанні практику майже не виникало труднощів, за вийнятком не очевидних позначеннь, та великої кількості необхідних експериментів в CoolPinkProgramm Приклад виводу після виконання скрипта:

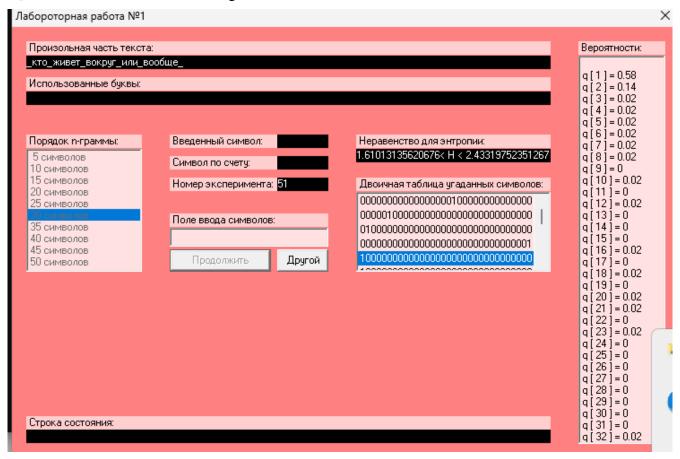
Таким чином, в виводі надані необхідні дані, такі як : H1 та H2, оцінка надлишковості (Redurancy) Приклад виконання CoolPinkProgramm для 10 символів:



Приклад виконання CoolPinkProgramm лдя 20 символів:



Приклад виконання CoolPinkProgramm для 30 символів:



Таблиця частот бувк:

к: 0.0329	ж: 0.0111
д: 0.0317	й: 0.0101
y: 0.0288	ш: 0.0097
п: 0.0268	x: 0.0075
ь: 0.0210	ю: 0.0066
я: 0.0205	э: 0.0036
г: 0.0192	ц: 0.0036
ч: 0.0179	щ: 0.0029
б: 0.0179	ф: 0.0018
ы: 0.0170	
з: 0.0154	
	д: 0.0317 у: 0.0288 п: 0.0268 ь: 0.0210 я: 0.0205 г: 0.0192 ч: 0.0179 б: 0.0179 ы: 0.0170

Таблиця частот для біграм:

('t', 'o'): 0.0180	('k', 'o'): 0.0095	('p', 'a'): 0.0077
('н', 'е'): 0.0126	('o', 'c'): 0.0095	('e', 'c'): 0.0076
('c', 't'): 0.0122	('а', 'л'): 0.0094	('л', 'и'): 0.0075
('o', 'b'): 0.0121	('o', 'h'): 0.0092	('и', 'т'): 0.0074
('н', 'о'): 0.0119	('\pi', 'p'): 0.0084	('e', 'T'): 0.0074
('н', 'а'): 0.0110	('p', 'o'): 0.0083	('T', 'b'): 0.0073
('e', 'h'): 0.0104	('B', 'o'): 0.0082	('o', 'p'): 0.0072
('п', 'о'): 0.0104	('k', 'a'): 0.0082	('o', 'm'): 0.0072
('r', 'o'): 0.0104	('н', 'и'): 0.0079	('e', 'p'): 0.0071
('o', 'T'): 0.0100	('a', 'k'): 0.0077	('t', 'e'): 0.0071

Таблиця частот біграм з кроком:

(- -), 0.0190	(-), 0,0004	(1-1 1-1), 0.0056
('o', 'o'): 0.0180	('и', 'о'): 0.0084	('e', 'ь'): 0.0056
('e', 'o'): 0.0129	('a', 'a'): 0.0081	('c', 'a'): 0.0055
('o', 'e'): 0.0114	('и', 'e'): 0.0077	('ч', 'о'): 0.0055
('a', 'o'): 0.0102	('c', 'o'): 0.0076	('e', 't'): 0.0055
('о', 'и'): 0.0099	('o', 'h'): 0.0075	('т', 'л'): 0.0052
('e', 'e'): 0.0097	('e', 'a'): 0.0073	('a', 't'): 0.0052
('o', 'a'): 0.0093	('o', 'T'): 0.0073	('B', 'p'): 0.0050
('и', 'a'): 0.0090	('e', 'h'): 0.0072	('и', 'и'): 0.0050
('a', 'e'): 0.0090	('e', 'и'): 0.0069	('л', 'н'): 0.0049
('a', 'и'): 0.0087	('B', 'e'): 0.0067	('o', 'c'): 0.0049

Таблиця частот біграм з пробілом:

(' ', ' '): 0.0260	(' ', 'и'): 0.0119	('r', 'o'): 0.0083
('o', ' '): 0.0235	('ь', ' '): 0.0108	(' ', 'k'): 0.0078
('e', ' '): 0.0187	('я', ' '): 0.0107	('м', ' '): 0.0077
('и', ' '): 0.0173	(' ', 'o'): 0.0105	('a', 'π'): 0.0074
(' ', 'B'): 0.0163	('н', 'е'): 0.0101	('k', 'o'): 0.0074
('a', ' '): 0.0157	('c', 'T'): 0.0096	('o', 'B'): 0.0073
(' ', 'н'): 0.0156	(' ', 't'): 0.0094	(' ', 'д'): 0.0072
(' ', 'c'): 0.0154	('н', 'о'): 0.0094	('y', ' '): 0.0071
(' ', '\pi'): 0.0152	('н', 'а'): 0.0088	('л', ' '): 0.0070
('t', 'o'): 0.0142	('п', 'о'): 0.0084	('e', 'h'): 0.0070

Таблиця біграм з пробілом та кроком:

(' ', ' '): 0.0404	('o', 'o'): 0.0116	(' ', 'ы'): 0.0063
(' ', 'o'): 0.0271	('e', ' '): 0.0116	(' ', 'и'): 0.0063
(' ', 'a'): 0.0221	(' ', 'p'): 0.0095	('e', 'e'): 0.0057
(' ', 'e'): 0.0217	('л', ' '): 0.0084	('a', 'o'): 0.0057
('T', ' '): 0.0176	('о', 'и'): 0.0075	('r', ' '): 0.0057
('н', ' '): 0.0173	('c', ' '): 0.0074	(' ', 'c'): 0.0057
('o', ' '): 0.0143	('e', 'o'): 0.0073	('o', 'h'): 0.0054
('a', ' '): 0.0121	('a', 'и'): 0.0067	('д', ' '): 0.0054
('и', ' '): 0.0116	(' ', 'н'): 0.0066	('c', 'o'): 0.0053
(' ', 't'): 0.0116	('o', 'e'): 0.0065	('e', 'и'): 0.0053

Висновки:

В даній лаборіторній роботі я навчився обраховувати ентропію відкритого тексту та виродженість мови, також екпериментальним шляхом я визначив частоти літер та різних біграм в доволі великому тексті, таким чином, можна стверджувати, що такі дані будуть актуальними для усієї мови