

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»**

ФІЗИКО- ТЕХНІЧНИЙ ІНСТИТУТ

Кафедра інформаційної безпеки

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

з дисципліни

Криптографія

З теми: « Експериментальна оцінка ентропії на символ джерела відкритого
тексту »

Перевірила:

Селюх П.В

Виконав студент групи
ФБ-94

Рябко Дмитро

Мета роботи:

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

Порядок виконання роботи

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H_1 та H_2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H_1 та H_2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H_1 та H_2 на тому ж тексті, в якому вилучено всі пробіли.
2. За допомогою програми CoolPinkProgram оцінити значення $(10) H$, $(20) H$, $(30) H$.
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Мета роботи:

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

Порядок виконання роботи

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H_1 та H_2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H_1 та H_2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H_1 та H_2 на тому ж тексті, в якому вилучено всі пробіли.
2. За допомогою програми CoolPinkProgram оцінити значення $(10) H$, $(20) H$, $(30) H$.
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Результати знаходяться в файлі Cryptolab1.xlsx

	Ентропія	R
Букви з пробілом	4,373672	0,125266
Букви без пробілу	4,4533	0,101105
Біграма з пробілом(перетин)	3,976199	0,999708
Біграма без пробілу(перетин)	4,144876	0,99623
Біграма з пробілом(без перетину)	3,982851	0,999801
Біграма без пробілу(без перетину)	4,148717	0,999625

H10

Лабораторная работа №1

Произвольная часть текста:
бы_смысла_говорить_что_футбольный_игрок_допустил_нарушение_если_бы_не_сущ

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: a

Символ по счету: 1

Номер эксперимента: 50

Неравенство для энтропии:
3,2628985693429 < H < 3,77406030335296

Двоичная таблица угаданных символов:

01000000000000000000000000000000	^
10000000000000000000000000000000	
00000000000100000000000000000000	
00000000000000000001000000000000	
00001000000000000000000000000000	
00000000000000000000000000000000	

Вероятности:

q[1] = 0,26
q[2] = 0,12
q[3] = 0,02
q[4] = 0,06
q[5] = 0,1
q[6] = 0,02
q[7] = 0
q[8] = 0,02
q[9] = 0,02
q[10] = 0
q[11] = 0,04
q[12] = 0,02
q[13] = 0,02
q[14] = 0
q[15] = 0,04
q[16] = 0
q[17] = 0,06
q[18] = 0,02
q[19] = 0,02
q[20] = 0,02
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0,02
q[27] = 0,02
q[28] = 0,02
q[29] = 0
q[30] = 0,02
q[31] = 0
q[32] = 0,06

Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

H20

Лабораторная работа №1

Произвольная часть текста:
туго_с_деньгами_а_то_что_вы_обещали_сделать_для_такого_то_старого_своего_пр

Использованные буквы:
й, ц, у, к, е, н, г, ш, щ, з, ф, ы, в, а, п, р.

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: о

Символ по счету: 17

Номер эксперимента: 50

Неравенство для энтропии:
 $4,46187300905082 < H < 3,9859791299831$

Двоичная таблица угаданных символов:

00000010000000000000000000000000
10000000000000000000000000000000
000000000000000000000000000000010000
100000000000000000000000000000000000
100000000000000000000000000000000000

Поле ввода символов:
о

Продолжить Другой

Вероятности:

q[1] = 0,18
q[2] = 0,04
q[3] = 0
q[4] = 0
q[5] = 0,04
q[6] = 0
q[7] = 0,02
q[8] = 0
q[9] = 0
q[10] = 0,04
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0,02
q[15] = 0,08
q[16] = 0,02
q[17] = 0,08
q[18] = 0,08
q[19] = 0
q[20] = 0
q[21] = 0,02
q[22] = 0,02
q[23] = 0,02
q[24] = 0,04
q[25] = 0,04
q[26] = 0,1
q[27] = 0,04
q[28] = 0,06
q[29] = 0,02
q[30] = 0
q[31] = 0,04
q[32] = 0

Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

H30

Лабораторная работа №1

Произвольная часть текста:
му_тогда_мы_так_ревностно_оправдываем_свое_не_совсем_порядочное_поведение_п

Использованные буквы:
й, ц, у, к, е, н, г, ш, щ, з, х, ф, ы, в.

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: а

Символ по счету: 15

Номер эксперимента: 50

Неравенство для энтропии:
 $5,53154623093661 < H < 3,97334044106812$

Двоичная таблица угаданных символов:

00000000000001000000000000000000
0000000000000000000100000000000000
0000000000000000000000000000000001
000000000000000000000000000000000001
000000000000000000000000000000000000

Поле ввода символов:
а

Продолжить Другой

Вероятности:

q[1] = 0
q[2] = 0
q[3] = 0,02
q[4] = 0
q[5] = 0,02
q[6] = 0,06
q[7] = 0,02
q[8] = 0,04
q[9] = 0
q[10] = 0,02
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0,08
q[15] = 0,08
q[16] = 0,02
q[17] = 0
q[18] = 0,1
q[19] = 0,04
q[20] = 0,06
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0,02
q[25] = 0,04
q[26] = 0,06
q[27] = 0,08
q[28] = 0,08
q[29] = 0,02
q[30] = 0
q[31] = 0
q[32] = 0,14

Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

$$3,262 < R(10) < 3,77$$

4,461<R(20)<3,98

5,531<R(30)<3,973