Міністерство освіти і науки України Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського» Фізико-технічний інститут

Лабораторна робота №1
З предмету «Криптографія»
На тему «Експериментальна оцінка ентропії на символ джерела відкритого тексту»

Виконали:

Студенти групи ФБ-83 Жоглик О.

Купрієнко А.

Перевірив:

Чорний О.М

Мета:

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 Н1 та Н2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення Н1 та Н2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення Н1 та Н2 на тому ж тексті, в якому вилучено всі пробіли.
- 2. За допомогою програми CoolPinkProgram оцінити значення (10) $\rm H$, (20) $\rm H$, (30) $\rm H$.
- 3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Хід роботи:

Таблиця частот монограм без пробілів

o: 0.114238
e: 0.0880262
a: 0.0757128
и: 0.0669517
н : 0.0634338
т: 0.0626406
c: 0.0542988
л: 0.0483407
в: 0.046071
p: 0.0440585
м: 0.0331322
д: 0.032174
к: 0.0316928
y: 0.0268882
п: 0.0262489
я: 0.0199967
ы: 0.0194812
ь: 0.0194138
б : 0.0184804
г: 0.0182701
3:0.016153
ч : 0.0136276
й: 0.0133925
ж: 0.0111119
ш: 0.0091254
x: 0.00875973
ю: 0.00649418
э: 0.00437299
щ: 0.00340656
ц: 0.0029694
ф: 0.000899068
ъ: 0.000137472

Таблиця частот монограм з пробілами

: 0.157488
o: 0.0962469
e: 0.0741631
a: 0.063789
и: 0.0564076
н : 0.0534437
т: 0.0527754
c: 0.0457474
л: 0.0407276
в: 0.0388154
p: 0.0371198
м: 0.0279143
д: 0.027107
к: 0.0267016
y: 0.0226536
п: 0.022115
я: 0.0168475
ь: 0.0164722
ы: 0.0164131
6: 0.0155699
г: 0.0153927
3: 0.0136091
ч : 0.0114814
й : 0.0112834
ж: 0.00936189
ш: 0.00768826
x: 0.00738018
ю: 0.00547143
э: 0.0036843
щ: 0.00287007
ц: 0.00250175
ф: 0.000757476

Таблиця частот біграм без пробілів з перетинами

C T

ПР

 $6\,0.00105029\,0.00026945\,0.00008523\,0.00001925\,0.00008523\,0.00001925\,0.00002887\,0.00257073\,0.00000825\,0.0001237\,0.00007561\,0.00000825\,0.0001237\,0.000001237\,0.000012$ $\textbf{B}\ 0.00716230\ 0.00017871\ 0.00040417\ 0.00017871\ 0.00040417\ 0.00017871\ 0.00040417\ 0.00017976\ 0.00005774\ 0.0002171\ 0.000057956\ 0.00001100\ 0.00001797$ $\pi \, 0.00530917 \, 0.00007698 \, 0.00103791 \, 0.00007698 \, 0.00103791 \, 0.000027219 \, 0.00015809 \, 0.00025845 \, 0.000025845 \, 0.00005224 \, 0.00015809 \, 0.00052845 \, 0.00000000 \, 0.00015809$ $e\,0.00019246\,0.00281818\,0.00412691\,0.00419977\,0.00543840\,0.00253499\,0.00130599\,0.00130$ $\pi\,0.00040829\,0.00174040\,0.00493525\,0.00082896\,0.00224355\,0.00082896\,0.00224355\,0.00082896\,0.00224355\,0.00082896\,0.00230266\,0.000150532\,0.00226004\,0.00150532\,0.00230266\,0.00089494\,0.00174040\,0.00085508\,0.00030244\,0.00000000\,0.00000000\,0.00000000\,0.000043441\,0.00049078\,0.00229304\,0.000150532\,0$ $\mathbf{m}\,0.00247587\,0.00037942\,0.00108878\,0.00027632\,0.00054027\,0.000475104\,0.00013335\,0.00026120\,0.00431388\,0.00000000\,0.00092244\,0.00004281\,0.00026759\,0.00037667\,0.000131286\,0.00004471\,0.00005886\,0.00041517\,0.00005774\,0.00002474\,0.00000000\,0.00104754\,0.00005636\,0.0003794\,0.00001925\,0.00078084$ $\\ o\ 0.00015672\ 0.00615463\ 0.00959556\ 0.00597454\ 0.00660141\ 0.00363751\ 0.00268208\ 0.00253636\ 0.00109703\ 0.000111902\ 0.00012098\ 0.00253636\ 0.00109703\ 0.00031619\ 0.00000000\ 0.00000137\ 0.00000000\ 0.00004967\ 0.0004969\ 0.00124825\ 0.00196310\ 0.00053636\ 0.00053636\ 0.000536000\ 0.00053600\ 0.0000000\ 0.0000000\ 0.0000000\ 0.000000\ 0.000000\ 0.0000000\ 0.0000$ $n\,0.00833494\,0.00016497\,0.00040142\,0.00021996\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087295\,0.00087296\,0.00002545\,0.00017047\,0.00009348\,0.00017047\,0.0009348\,0.000009348\,0.000009348\,0.00009348\,0.00009348\,0.000009348\,0.00009348\,0.00009348\,0.00009348\,0.00009348\,0.00009348\,0.00009348\,0.00009348\,0.00009348$ $\ \ \, \tau\,0.00615875\,\,0.00029144\,\,0.00496962\,\,0.00013197\,\,0.00037805\,\,0.00069352\,\,0.00009073\,\,0.00014847\,\,0.00001379\,\,0.00014847\,\,0.00001237\,\,0.00014847\,\,0.00004812\,\,0.000$

H O

и й к л м

Таблиця частот біграм без пробілів без перетинів

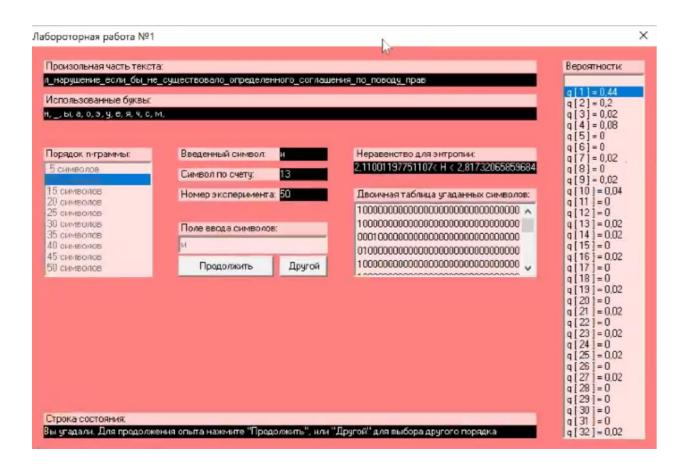
 $a\ 0.00023095\ 0.00151219\ 0.00490776\ 0.00162217\ 0.00276869\ 0.00162217\ 0.00276869\ 0.00162217\ 0.00276869\ 0.00130599\ 0$ $\begin{array}{l} \begin{array}{l} \beg$ $\pi \ 0.00849578 \ 0.00035193 \ 0.00055814 \ 0.00032718 \ 0.0000275 \ 0.00000000 \ 0.0005814 \ 0.0000275 \ 0.00000000 \ 0.00062137 \ 0.0002749 \ 0.0000275 \ 0.00000000 \ 0.00062137 \ 0.000275 \ 0.00000000 \ 0.00062137 \ 0.0008248 \ 0.00119601 \ 0.00194936$ $\mathbf{m}\,0.00254323\,0.00038217\,0.00108328\,0.00028594\,0.000259420\,0.00043441\,0.00265871\,0.000132248\,0.00024470\,0.00043441\,0.00265871\,0.00035743\,0.00132248\,0.00002406\,0.00005499\,0.00002200\,0.00000000\,0.00106403\,0.00007149\,0.00007149\,0.00007149\,0.000073685$ $\pm 0.01072833\ 0.00025845\ 0.00039867\ 0.00025845\ 0.00039867\ 0.00025845\ 0.00039867\ 0.0002647\ 0.00027\ 0.0002647\ 0.00027\ 0.0002647\ 0.00027\$ $c\,0.00163867\,0.00021446\,0.00291716\,0.00021446\,0.00291716\,0.00011548\,0.00047290\,0.00373099\,0.00007149\,0.0003749\,0.0002749\,0.0002749\,0.0002749\,0.00007149\,0.000007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149\,0.00007149$ $\ \ \, \tau\,0.00618900\,\,0.00027494\,\,0.00500124\,\,0.00010998\,\,0.00037494\,\,0.00500124\,\,0.00010998\,\,0.00035743\,\,0.00069385\,\,0.00009348\,\,0.00015672\,\,0.00500949\,\,0.000015672\,\,0.00500949\,\,0.000015672\,\,0.00500949\,\,0.00001375\,\,0.00180363\,\,0.00783591\,\,0.00012922\,\,0.00016497\,\,0.00069836\,\,0.00001100\,\,0.00003849\,\,0.0001100\,\,0.00003849\,\,0.00011100\,\,0.00003849\,\,0.00011100\,\,0.00003849\,\,0.00011100\,\,0.00003849\,\,0.00011100\,\,0.00003849\,\,0.00011100\,\,0.00003849\,\,0.00011100\,\,0.000011100\,\,0.00003849\,\,0.00011100\,\,0.0000111000\,\,0.0000111000\,\,0.00000111000\,\,0.0000111000\,\,0.0000111000\,\,0.0000111000\,\,0.0000111000\,\,0.000011000\,\,0.0000110$ $\ \ \, \mathbf{x}_0 \ \, 0.00013747 \ \, 0.00058288 \ \, 0.00034368 \ \, 0.00008248 \ \, 0.00008248 \ \, 0.00008248 \ \, 0.00008248 \ \, 0.00008248 \ \, 0.00008248 \ \, 0.00008248 \ \, 0.00008249 \ \, 0.00008248 \ \, 0.00008249 \ \, 0.00008$

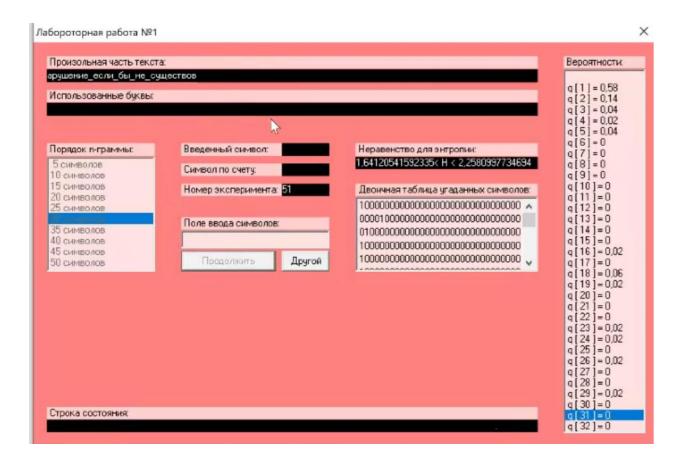
Таблиця частот біграм з пробілами з перетинами

 $\\ e 0.00003166\ 0.00160898\ 0.00172039\ 0.00322031\ 0.00383716\ 0.00172039\ 0.00322031\ 0.00383716\ 0.00176730\ 0.00077165\ 0.0001784891\ 0.000176730\ 0.00077165\ 0.000176730\ 0.00077165\ 0.000176730\ 0.00077165\ 0.000176730\ 0.00077165\ 0.0001767300\ 0.00017673$ $\mathbf{m}\,0.00174502\,0.00002345\,0.00006885\,0.00004808\,0.00000469\,0.0016770\,0.0000173\,0.0002345\,0.00000000\,0.000372576\,0.00000000\,0.0000173\,0.00023405\,0.00000000\,0.00001407\,0.00002815\,0.00000000\,0.00001407\,0.000082091\,0.00004808\,0.00000000\,0.00005269\,0.000889514$ $p\ 0.00697773\ 0.00013252\ 0.00030022\ 0.00017943\ 0.00070598\ 0.000695076\ 0.00017231\ 0.000070598\ 0.000695076\ 0.00017231\ 0.000070598\ 0.000695076\ 0.00017231\ 0.000070598\ 0.000695076\ 0.00017231\ 0.000070598\ 0.0000070598\ 0.000070599\ 0.000070599\ 0.000007059\ 0.000007059\ 0.000007$ $\mathbf{h} 1, 0,000000000,00013135,0,00078573,0,00008561,0,00010789,0,00116804,0,00000000,0,00016301,0,$ $0.00254130\ 0.00620725\ 0.01455941\ 0.00268437\ 0.00667282\ 0.00749842\ 0.0045297\ 0.00268437\ 0.00667282\ 0.00033657\ 0.00002345\ 0.0020295\ 0.000033657\ 0.00002345\ 0.0020295\ 0.00402597\ 0.0028462\ 0.00033657\ 0.0000033657\ 0.0000033657\ 0.0000033657\ 0.0000033657\ 0.0000033657\ 0.0000033657\ 0.0000033657\ 0.000000000000000$

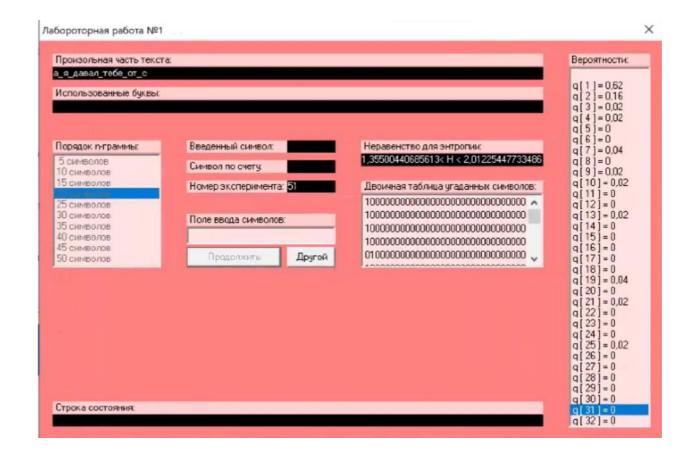
Таблиця частот біграм з пробілами без перетинів

 $6\,0.00087485\,0.00023220\,0.00006567\,0.00001642\,0.00002580\,0.0001545\,0.00002320\,0.00006567\,0.00001642\,0.00002580\,0.0012585\,0.00005864\,0.00105545\,0.00005864\,0.000005864\,0.000005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.000005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.00005864\,0.0000$ $\ \ \, \tau\,0.00507790\,\,0.00002111\,\,0.00372927\,\,0.00000000\,\,0.00042453\,\,0.00000000\,\,0.00008209\,\,0.00541799\,\,0.00000000\,\,0.000137209\,\,0.001267482\,\,0.0001335\,\,0.00245803\,\,0.001335\,\,0.00245803\,\,0.001335\,\,0.00245803\,\,0.001335\,\,0.00245803\,\,0.001335\,\,0.00245803\,\,0.001335\,\,0.00245803\,\,0.001335\,\,0.00245803\,\,0.001335\,\,0.00245803\,\,0.0013350\,\,0.00245803\,\,0$ $0.00242520\ 0.00614978\ 0.01464969\ 0.00269023\ 0.00655554\ 0.00751717\ 0.00465103\ 0.00026504\ 0.00115631\ 0.00044564\ 0.00432736\ 0.00056760\ 0.00010789\ 0.00032133\ 0.0003049\ 0.000349120\ 0.0003775\ 0.00159022\ 0.00290602$





	Ентропія	Надлишковість
Монограма(з пропусками)	4.386	0.122
Монограма(без пропусків)	4.461	0.108
Біграма з перет. та проп.	3.976	0.204
Біграма з перет та без проп.	4.138	0.172
Біграма без перет та з проп.	4.137	0.172
Біграма без перет та без проп	3.976	0.204
H(10)	2.110 <h<2.817< td=""><td>0.436<r<0.578< td=""></r<0.578<></td></h<2.817<>	0.436 <r<0.578< td=""></r<0.578<>
H(20)	1.355 <h<2.012< td=""><td>0.597<r<0.729< td=""></r<0.729<></td></h<2.012<>	0.597 <r<0.729< td=""></r<0.729<>
H(30)	1.641 <h<2.258< td=""><td>0.548 <r<0.671< td=""></r<0.671<></td></h<2.258<>	0.548 <r<0.671< td=""></r<0.671<>



Висновки:

Отже в даній лабораторній роботі ми дослідили ентропію для монограм та біграм, з пробілами так і без них, знаходили надлишковість російської мови. Найбільша складність виникала при роботі з форматуванням таблиць та фінального розрахунку ентропії