

КРИПТОГРАФІЯ
КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

5 варіант

ФБ-01 Пітель Богдан

ФБ-01 Ширий Віталій

Експериментальна оцінка ентропії на символ джерела відкритого тексту

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.

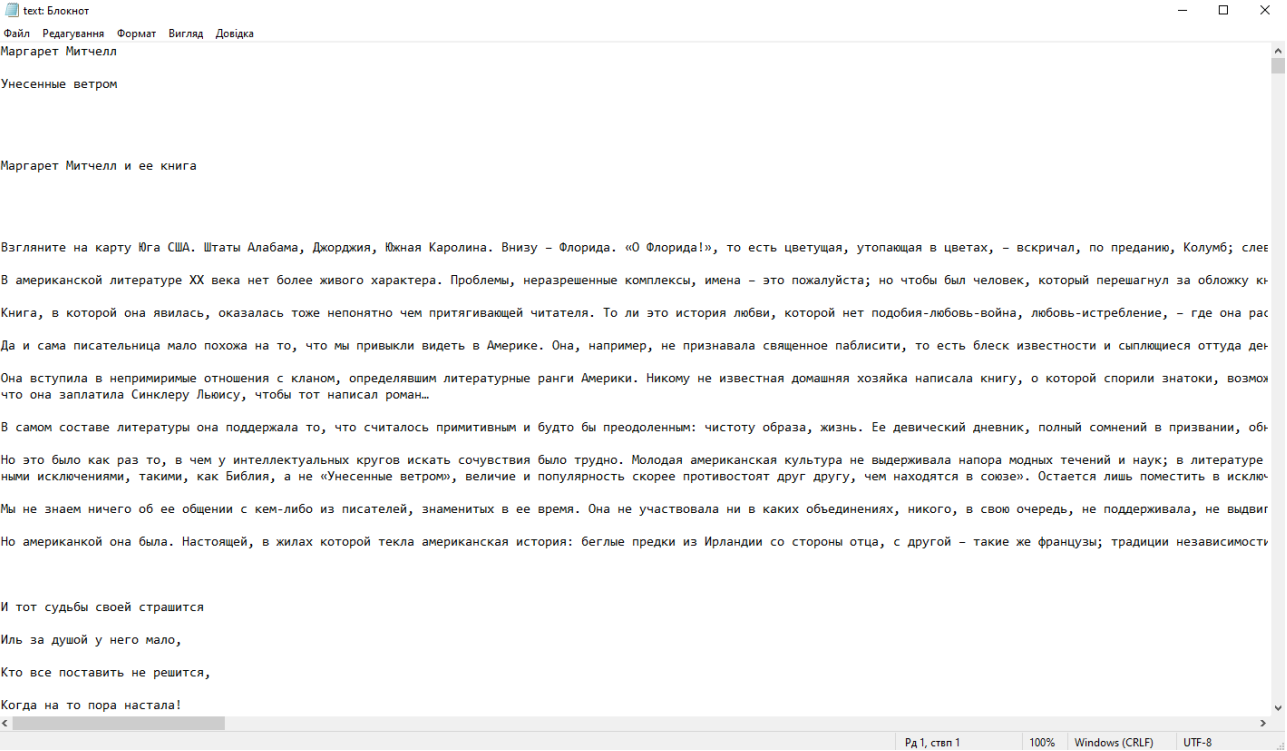
Порядок виконання роботи

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.

Хід роботи

1. Знаходження потрібного за розміром тексту (щонайменше 1МБ).
2. Текст «Унесенные ветром» (рос. м.)
3. Почистимо наш текст від символів, які відсутні у алфавіті.
4. Підрахуємо кількість частот букв і частот біграм в тексті.
5. Підрахуємо H_1 та H_2 , також частоти букв та біграм.
6. За допомогою програми CoolPinkProgram оцінимо значення $(10)H$, $(20)H$, $(30)H$.
7. Оцінимо надлишковість російської мови в різних джерелах.

1-3. Текст не відредагований (1.81 МБ)



Використасмо python код, для того, аби почистити текст від символів.

Збережемо цей файл, як withspace.txt.



У наступному кроці приберемо з тексту пробіли, і також збережемо, але під назвою withoutspace.txt

text_withoutspace: Блокнот

Файл Редагування Формат Вигляд Довідка

маргаретмитчеллунесенієветроммаргаретмитчеллієекнигазгягнятенакартуюасшаштатабаламаджорджияжжянакаролинавнзюфлоридаофлоридатоєстьцветущаутопащацветатхвскричалпопредани
рянавытравливаниесобеихсторонтолидамскйроманподнявшийсядонастоящейлитературыпотомучтотолькодаманаверномоглоподсмотретьзасвоейгероинейкактацелуетсебявзеркаленожествоводругихб
мяуступаюместодругдругуноигарантируякаждоуположениеивистериилитературыворимойнаглазахсоединеннымудромомассовыхсредствэтоткомбинатполучивдвдурвбестселлерыночередногокандидат
онавысказаласьтакеслиториякоторуюхочешьрассказатьихаактерыневыдерживаютпростотычтотоназываетсяголойпрозойлучшихоставитьидитбоганыстилистинемоглабымбытьеслибыихотеланозтыбы
какбиблияанеунесенієветромвеличиеипопулярностьскореепротивостоятдругдругученнаходятсясвоозеетасяслишьпоместитьвыисключениясервантесайданераблостотогочеховадикенсамаркатве
анцузьттрадицинезависимостиполаганиянасобственнуюсилуготовностьрисковатьлюбимыестихамиематерибылиитотсудьбысвоейстрашитсяиздадуйнегомалоковсепоставитьнерешитсягоданат
гообществаидимосемейныепреданияпечатленияностипривелиеекстранноймысличтоонаживетзавоеваннойстранекакимипутямбылозадуманнозвоеваниемынезнаемхарактербылскрытыйствавший
эподалекутпрославленногойифинансоваяристократиясменилаземельнуюевстранишнейописатисториипротиворечияпрогрессасказалисьсособойстротойхищничествовспекуляциичиничноеграф
дачинамириноеурегулированиеобалагеряомобилизовалисьгдоупрезидентомсталаубежденныйаболитониставраамликольнжнестытотделилисьобразовавконфедерациювоенныедействияначалисьперее
епостыхлхдипридумалиногоинишхидансоваяристократиясменилаземельнуюевстранишнейописатисториипротиворечияпрогрессасказалисьсособойстротойхищничествовспекуляциичиничноеграф
осказалисьнаобразекизинострантосдапришлотовдручающеесяочетанедикостицивилизацииминуякултуруетьпредставлениеосмыслектороевсегдактостанавлялодругуеамерикисталоувств
жностьслышатьсмыслделаствасабсолютнойпешойвраухпхотакателяскажитеуслышитсамомделездесьедьбылпроведенэтотопытжтьбезвластиебезправительствамечтаваамерикибылапровозглашеная
льностьискушениесильноушалиотнег описательницавполнесказатьтруднотакпокрайнеймеревпервомвпечатленииможнобылопонятьзаглавиееекнигивзятоизстихотворениягорацияпопулярномуае
азалсьещежестчэтопородачистосторнаментальнаянаистрочкаиизбалладмуратакспитгордияпрежнихднейтожекузчаланедаромдетигордоститакназвалисторикробертмэнсонмариерсвойтредхтомниквышедший
ивопоставляетисториюсвоегодядьшкпитеризвестномудядетомубичерстоукакрисуеткормилицуманушкобольшогоосамаидратинотыумитчеллсвоюзатдажеанмереновызывающегоданпримеронаначинае
сплескасословногоовискоммериябылобыльшимупрощениеммызнаемивеликиепримерыкактостойапредисловиквоиничемируобъявлялчтоегоинтересуютнизишесословияегоониколайростоводнимолоде
встаетсбавчужойдушкаквсобственнойвекразобщенияэтчертанезаменимасоскарлеттдобавилосьотчтоонапопалавположениеблизкоевновойамерикенотимчеловекбезкультурысильнымиистрымук
веванглийскомзвучающийкоторогостыиальицветомнаподобиенашегогоричиветаиболеезньсохраняетсякарлатинаибудницаназвербагряномнеподавалсяэтотхарактеримениямкс
нымчтмосиохарасталацетвореданациональнойгероинядумаючтоэтооченьсквернодляморальногуюмственногоосостояниянацииеслинацияспособнааплодироватьиувлекатьсяжндейнойоторавеласеб
нытвьвернепочемуэтомоложедолженажелатьобеспеченностипредоставьтэзотстарымустальменияизумляетнекоторыехныхкакэтоонинасколькомогулпнатьнепрототоскутпообеспеченностиюве
былотвергнутечаяносравниепредметсвоейстрастисоскарлеттонанаписаларадачтоаманравитсяскарлетттычтвысчитаетееерешительнойдстойнойосхищениядевушкойпомнеонабыладалеконевосж
ивостьэтомсмыслонабыласетройфолкнерахотиянепризнаннойушлииунесенієветромнаостсяидеалнеисполнимыйсонкоторыйвидитскарлеттдушлетящаявбелойтуманноймассепочтиголевскаястру
инимитчеллодаособеннопоразилаегочертылицаебылинежнытоникитотозатанеоеигрустноебылвеепрекрасноминенногооргомвзглядеонабылаонанемоглоанебытьвышесеезтойтольпнечастныхжени
новомнарядекоторыйтаклюбилиаврусскомязкечастьглаваскарлетттохаранебылакрасавицейномужичиньярдиотдавалисебезвзломтчетеслиониподоблизнецамтарлтонамстановилисьжертвामीеяро
чонисзеленымисафьяновымтуфелькамибезакбуквольтыкочтопривезенныеийецотомизатлантйиуфлатъкакнелъзаблөөегооднообъягивалбезупрекнутилибесспорноамуютконкувтрехграфствихатэ
истивбелоснежныерозыветущихкизильныхдеревьевмерцаликосыхлучахзакатногосолнцалошадиблизнецовкупныеживотныезолотистогнедшеподстатьшевельрамсвоиххозяевстолихуконозанаподье
емудростамиграфствоклөөлтонсевеириджорджибылоещемолодойжизнътаннавазляджителейчарльстонасаванимогастыпокачтонеуатриланекоторогоналетграубостиболесетарыеистепенныеобитатели
етауказавшегоимнадобойдвозвратилисьдомойместесидомоймешаебываеотаватьствястенаучебногозаведениядемлишисньекдворостяртибрентрасмотривалимолад
тоодинаразговоряшйилиукиетгоотцельконапрошлойнеделеговорилапечатнонашипредставителивавшийтонепридутктомусамомукободоприемлемомусоглашениисмистерииликольномоповодукофед
ичьчемлибодругомоченьрадачтоджорджиянездумалаотделатьсядо святокиначунабылибиспорченывсерождествовскиебалыиещеразуслышупровойнуидувдомиможнобылосомневатьсячтоонас
исегодняздаромануотропомкоаноанежелатитомзаселуфонтейновампоскакалисодзавчеравечеромгодавьявилисьдомойонатоженегонесказалавечеравечеронамповезлокакразпреднашпириездомп
русскакалипораньшепокаонавнашестеласяобойдоталсьееумасливатькаквдумаеетоназвудетбойдаскарлетткакивсежителиграфствапростонемоглаосвотитьсясымыслчтотокрошкисистарлтондерж
рбекуюилксамназойнойовлошадноапоехалабыдапапсказалчтотоопаснолошадьслишкомгорячанудевчонокийнедадутонизавиличтоонадолжнахотьябыразприехатьгостикакприличествуетдамеажик
еболегостымбарецонмвляжневывороченнеластыземлиналиновыенаподсыхающихребяихбороздиливатопунуовыеибурневгустойтенилежалиалкахлапоковызерепосевавыбеленныйизвесткойкирипич
нственнымчутьзловещимнаполненнмтерпеливымвекowymшорохомверхушкассосепохожимнавоздохлинугрозуберегисьберегистыужзарасталооднаждыполемыможемзавладетьтобосноводослухасиде
натьчтоовиприскачетдомойнемоглаеберечьтанццдляласапотомстатьсясносимподпиратьстенкуэтотыблизнецовлушительнорахоталисьвотчтомалыткатыдолжнадатынепервыйвалъсастпоп
омолакнзутодлямениченностьразочарованнопротнугласкарлеттзатотдурочкочарлигамилътонееплмниикобручтсясмилочкойуилксвенужедавныдавиновзвестночтоонидолженженитьсяхотлониче
екретитытеперьдолжнапообещатьчтосдаешужинатьснаминуконачесваимашинальнонопроботмаласкарлетттибеаешьотдатынамсвальсьобещакнутыпрелестьюображажакксамальчишказбеситсялау
иваядругдругандовольнопрозрачнонамекачтотприближаетсявремяужинамолчанискарлеттнесразушлодоихсознанияоназавсезовремянепроронилисловаонакоңионцошутиликакуютоперемнусия
иваякрасунопольизамеривожиданияхозяевюноириспрошалисьпообещавскарлеттприехатьзавтракуилксавпронаншихдятьеатматембежалискрльцавсочилисьдланоспровождаетмеджисомустили
илинонепридалэтомужначениячтозотсейкактыдумашьнепоймужетмысказаличтонибудткакчетоерассердилонаминутубапогузилисьвразмышленияничеготаконегомугрипоминитьпритомкогдакае

4. Підрахуємо кількість літер у тексті, частоти букв біграм.

Кількість літер:

Counter({'o': 86795, 'e': 71488, 'a': 68880, 'и': 55334, 'н': 54027, 'т': 51428, 'с': 43763, 'л': 42313, 'р': 34703, 'в': 34532, 'к': 28029, 'д': 25794, 'м': 25774, 'у': 23104, 'п': 22345, 'ы': 15956, 'ь': 15887, 'я': 15788, 'б': 13891, 'г': 13813, 'э': 13284, 'ч': 11408, 'ж': 9577, 'й': 9522, 'х': 8040, 'ш': 7486, 'ю': 4636, 'э': 3577, 'ц': 2884, 'щ': 2726, 'ф': 910, 'ь': 158})

Частота, для зручності додали список у letter_frequencies.xlsx

letter	frequency
м	0,031514
а	0,084221
р	0,042432
г	0,016889
е	0,087409
т	0,062882
и	0,067658
ч	0,013949
л	0,051737
у	0,02825
н	0,06606
с	0,05351
ы	0,01951
в	0,04223
о	0,106126
к	0,034271
э	0,016243
я	0,019304
ю	0,005669
ш	0,009153
б	0,016985
д	0,031539
ж	0,01171
ф	0,001113
ь	0,019425
ц	0,003526
щ	0,003333
п	0,027322
х	0,009831
й	0,011643
э	0,004374
ъ	0,000193

Freqbigramwithspace:

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я	
а	0.001847	0.000876	0.014883	0.003753	0.000828	0.004717	0	0.00214	0	0.009918	9.21e-08	0.000384	0.002635	0.007432	0.013934	0.01127	0.019394	0.000399	0.000481	0.007305	0.004821	0.000478	0.000323	0.000438	0.000848	0.000014	0	0	0	0.001003	0.000078	0.002127	
б	0.0019374	1.19e-06	0.000702	0.003083	0.000311	0.002184	0.000833	0	0.001132	0	0.000119	0.000491	0.001482	0.000714	0.000335	0.000249	1.43e-08	0.000388	0.000385	0.000303	9.31e-03	9.92e-05	0.001213	0.000384	0.000713	0.000073	0.000037	0.000012	0	0	0.000146	0.000778	0.002035
в	0.000213	0.000871	2.05e-06	5.43e-05	0.07e-06	1.23e-05	0.000239	1.73e-05	0	0.000218	0	0.000131	0.000753	8.93e-05	0.000214	0.000339	0	0.001144	0.000141	1.15e-05	0.000219	0	0.000143	0.000148	0.000143	0.000148	0.000143	0.000148	0.000143	0.000148	0.000143	0.000148	
г	0.000647	0.000278	0.07e-06	3.99e-05	0.21e-05	0.000283	0.000466	0	0	0.000237	0	0.00013	0.000176	0.000714	0.000011	0.000934	0.000261	0.000013	0.000038	0.000153	0.000681	0	0.133e-05	0.99e-05	0.000128	1.02e-06	0	0.002481	0.000169	0	0	0.000238	
д	0.000083	0.001022	0	1.02e-06	1.02e-06	0.001245	0.000133	0	0	0.000838	0	0.000139	0.000774	1.02e-06	0.000284	0.000483	0	0.000441	1.13e-05	0.17e-05	0.00063	0	0	0.98e-05	4.09e-06	0	0	0	0	0	0	0	
е	0.001489	0.000789	5.93e-05	0.000026	1.34e-05	0.39e-05	0.000121	0	0.000241	0	0.000024	0.000129	0.000552	4.4e-05	0.000409	0.000145	0.000425	0.000094	0.000094	0.000201	1.14e-05	0.000174	2.05e-06	0.77e-05	0.000128	2.98e-05	0	2.84e-05	0.000873	0.000727	9.21e-06	0.000143	
ж	0.000213	1.27e-05	0.001245	0.000148	0.000781	0.002184	0.000213	0	0.000762	0	0.000132	0.000287	0.000131	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133	0.000133		
з	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
и	0.00025	0.000173	1.74e-05	2.05e-06	0.39e-05	0.000133	0.000243	0	2.78e-05	0	0.001439	0	9.92e-05	2.26e-05	1.44e-06	0.000893	0.000282	0	2.05e-06	2.15e-05	1.12e-06	0.000133	0	0	0.102e-06	0.000021	0	0	0	0	0.45e-05	0	
й	0.0189	1.12e-05	0.00021	0.000292	0.000483	0.0004	0.00189	0	0.001893	0	0.0004	0.000213	0.000127	0.000483	0.000478	0.000705	0.000177	0.000128	0.000478	0.000566	7.47e-05	7.78e-05	0.000071	0.000132	0.00008	0.000269	0	0	1.84e-05	0.000125	0.000723		
к	5.42e-05	0.00127	0	0	0	0	0.001897	0	0	0.00131	0.000174	0.000779	0	0.00013	0.000174	0.000779	0	0.00013	0.000174	0.000779	0	0	0	1.16e-06	2.86e-05	4.09e-06	1.02e-06	0	0.000423	3.98e-05	7.77e-05	0.000196	
л	0.000399	0.007118	8.11e-05	0.000319	0.00018	0.000486	0.00154	0	0.000393	0	0.000218	0	0.00039	0.001898	0.000778	0.000779	0.000834	8.62e-05	0.000114	0.000099	0.000193	1.13e-05	0.000415	2.15e-05	7.98e-05	0.000137	1.64e-05	0	0.000128	0.000434	3.95e-05	9.31e-05	
м	0.000444	0.000219	0	9.31e-05	1.02e-06	0	0.000094	0	1.24e-05	0	0.000309	0	0.000459	0.000017	1.91e-05	0.000248	0.000249	0.000171	0.000072	0.000081	0.000093	1.11e-05	0.000028	4.71e-05	0.000031	9.10e-05	0	1.33e-05	0.000314	0.000483	0.000018	0.000125	
н	0.000307	0.000889	5.01e-05	0.000389	1.97e-06	0.00004	0.000796	0	0.000308	0	0.000795	0	0.000459	0.000584	6.83e-05	0.000388	0.000867	7.26e-05	0	0.000718	0.000217	0.000128	4.09e-06	3.07e-06	0	0.000041	2.05e-06	0	0.000837	0.000337	2.05e-05	0.000891	
о	0.000725	0.002878	1.93e-05	7.16e-06	0.86e-05	0	0.000809	0	0.000849	0	0.000101	0.000221	3.38e-05	0.001134	0.000212	0.000103	0.000103	0.000119	0.000132	0.000131	0.000131	7.16e-06	2.05e-06	1.13e-05	2.33e-05	0	2.05e-06	0.000911	7.06e-05	4.9e-05	1.02e-06	0.000173	
п	0.000278	0.010595	2.05e-06	1.33e-05	0.000141	0.000452	0.000094	0	1.64e-05	0	0.000776	0	0.000464	3.07e-06	0.000129	0.000285	0.000885	0	0.000134	0.000461	0.0002	0.002478	0.000157	4.09e-06	0.000436	0.000018	6.14e-06	0.000395	0.000026	0	0.000133	0.001187	
р	0.0189	1.12e-05	0.00021	0.000292	0.000483	0.0004	0.00189	0	0.001893	0	0.0004	0.000213	0.000127	0.000483	0.000478	0.000705	0.000177	0.000128	0.000478	0.000566	7.47e-05	7.78e-05	0.000071	0.000132	0.00008	0.000269	0	0	1.84e-05	0.000125	0.000723		
с	5.42e-05	0.00127	0	0	0	0	0.001897	0	0	0.00131	0.000174	0.000779	0	0.00013	0.000174	0.000779	0	0.00013	0.000174	0.000779	0	0	0	1.16e-06	2.86e-05	4.09e-06	1.02e-06	0	0.000423	3.98e-05	7.77e-05	0.000196	
т	0.000399	0.007118	8.11e-05	0.000319	0.00018	0.000486	0.00154	0	0.000393	0	0.000218	0	0.00039	0.001898	0.000778	0.000779	0.000834	8.62e-05	0.000114	0.000099	0.000193	1.13e-05	0.000415	2.15e-05	7.98e-05	0.000137	1.64e-05	0	0.000128	0.000434	3.95e-05	9.31e-05	
у	0.000444	0.000219	0	9.31e-05	1.02e-06	0	0.000094	0	1.24e-05	0	0.000309	0	0.000459	0.000017	1.91e-05	0.000248	0.000249	0.000171	0.000072	0.000081	0.000093	1.11e-05	0.000028	4.71e-05	0.000031	9.10e-05	0	1.33e-05	0.000314	0.000483	0.000018	0.000125	
ф	0.000307	0.000889	5.01e-05	0.000389	1.97e-06	0.00004	0.000796	0	0.000308	0	0.000795	0	0.000459	0.000584	6.83e-05	0.000388	0.000867	7.26e-05	0	0.000718	0.000217	0.000128	4.09e-06	3.07e-06	0	0.000041	2.05e-06	0	0.000837	0.000337	2.05e-05	0.000891	
х	0.000725	0.002878	1.93e-05	7.16e-06	0.86e-05	0	0.000809	0	0.000849	0	0.000101	0.000221	3.38e-05	0.001134	0.000212	0.000103	0.000103	0.000119	0.000132	0.000131	0.000131	7.16e-06	2.05e-06	1.13e-05	2.33e-05	0	2.05e-06	0.000911	7.06e-05	4.9e-05	1.02e-06	0.000173	
ц	0.000278	0.010595	2.05e-06	1.33e-05	0.000141	0.000452	0.000094	0	1.64e-05	0	0.000776	0	0.000464	3.07e-06	0.000129	0.000285	0.000885	0	0.000134	0.000461	0.0002	0.002478	0.000157	4.09e-06	0.000436	0.000018	6.14e-06	0.000395	0.000026	0	0.000133	0.001187	
ч	0.0189	1.12e-05	0.00021	0.000292	0.000483	0.0004	0.00189	0	0.001893	0	0.0004	0.000213	0.000127	0.000483	0.000478	0.000705	0.000177	0.000128	0.000478	0.000566	7.47e-05	7.78e-05	0.000071	0.000132	0.00008	0.000269	0	0	1.84e-05	0.000125	0.000723		
щ	5.42e-05	0.00127	0	0	0	0	0.001897	0	0	0.00131	0.000174	0.000779	0	0.00013	0.000174	0.000779	0	0.00013	0.000174	0.000779	0	0	0	1.16e-06	2.86e-05	4.09e-06	1.02e-06	0	0.000423	3.98e-05	7.77e-05	0.000196	
ъ	0.000399	0.007118	8.11e-05	0.000319	0.00018	0.000486	0.00154	0	0.000393	0	0.000218	0	0.00039	0.001898	0.000778	0.000779	0.000834	8.62e-05	0.000114	0.000099	0.000193	1.13e-05	0.000415	2.15e-05	7.98e-05	0.000137	1.64e-05	0	0.000128	0.000434	3.95e-05	9.31e-05	
ы	0.000444	0.000219	0	9.31e-05	1.02e-06	0	0.000094	0	1.24e-05	0	0.000309	0	0.000459	0.000017	1.91e-05	0.000248	0.000249	0.000171	0.000072	0.000081	0.000093	1.11e-05	0.000028	4.71e-05	0.000031	9.10e-05	0	1.33e-05	0.000314	0.000483	0.000018	0.000125	
ь	0.000307	0.000889	5.01e-05	0.000389	1.97e-06	0.00004	0.000796	0	0.000308	0	0.000795	0	0.000459	0.000584	6.83e-05	0.000388	0.000867	7.26e-05	0	0.000718	0.000217	0.000128	4.09e-06	3.07e-06	0	0.000041	2.05e-06	0	0.000837	0.000337	2.05e-05	0.000891	
э	0.000725	0.002878	1.93e-05	7.16e-06	0.86e-05	0	0.000809	0	0.000849	0	0.000101	0.000221	3.38e-05	0.001134	0.000212	0.000103	0.000103	0.000119	0.000132	0.000131	0.000131	7.16e-06	2.05e-06	1.13e-05	2.33e-05	0	2.05e-06	0.000911	7.06e-05	4.9e-05	1.02e-06	0.000173	
ю	0.000278	0.010595	2.05e-06	1.33e-05	0.000141	0.000452	0.000094	0	1.64e-05	0	0.000776	0	0.000464	3.07e-06	0.000129	0.000285	0.000885	0	0.000134	0.000461	0.0002	0.002478	0.000157	4.09e-06	0.000436	0.000018	6.14e-06	0.000395	0.000026	0	0.000133	0.001187	
я	0.0189	1.12e-05	0.00021	0.000292	0.000483	0.0004	0.00189	0	0.001893	0	0.0004	0.000213	0.000127	0.000483	0.000478	0.000705	0.000177	0.000128	0.000478	0.000566	7.47e-05	7.78e-05	0.000071	0.000132	0.00008	0.000269	0	0	1.84e-05	0.000125	0.000723		
а	0.001847	0.000876	0.014883	0.003753	0.000828	0.004717	0	0.00214	0	0.009918	9.21e-08	0.000384	0.002635	0.007432	0.013934	0.01127	0.019394	0.000399	0.000481	0.007305	0.004821	0.000478	0.000323	0.000438	0.000848	0.000014	0	0	0.001003	0.000078	0.002127		
б	0.0019374	1.19e-06	0.000702	0.003083	0.000311	0.002184	0.000833	0	0.001132																								

	With space	Without space
H1	4.376164965459297	4.462591812647137
H2	7.955477072143008	8.30571926842366

H(10)

Лабораторная работа №1

Произвольная часть текста:
ли_не_пра

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:
Символ по счету:
Номер эксперимента: 51

Неравенство для энтропии:
 $2,55040208003973 < H < 3,193789622077$

Двоичная таблица угаданных символов:
10000000000000000000000000000000
10000000000000000000000000000000
00000000000000000000000000000001
01000000000000000000000000000000
00000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000

Поле ввода символов:
Продолжить Другой

Строка состояния:

$$2.55 < H < 3.193$$

H(20)

Лабораторная работа №1



Произвольная часть текста:
какой_нибудь_страны

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:
Символ по счету:
Номер эксперимента: 51
Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
2,04436573750085 < H < 2,721928094887
Двоичная таблица угаданных символов:
00000100000000000000000000000000
00000000000000000000000000000100000
100000000000000000000000000000000000
00000000000000000000000000000000100
0000100000000000000000000000000000
0000000000000000000100000000000000

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

Строка состояния:

$2.04 < H < 2.72$

H(30)

Лабораторная работа №1

Произвольная часть текста:
_этом_конечно_не_имелось_в_ви

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:

Символ по счету:

Номер эксперимента: 51

Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
 $2,58773104502519 < H < 3,268434573446$

Двоичная таблица угаданных символов:

00000000000000000000000000000000	10000000000000000000000000000000
00000000000000000000000000000000	10000000000000000000000000000000
00000000000000000000000000000000	10000000000000000000000000000000
10000000000000000000000000000000	00000000000000000000000000000000
00000100000000000000000000000000	00000000000000000000000000000000
10000000000000000000000000000000	00000000000000000000000000000000

Вероятности:

$q[1] = 0,34$
$q[2] = 0,18$
$q[3] = 0,08$
$q[4] = 0,06$
$q[5] = 0,02$
$q[6] = 0,04$
$q[7] = 0$
$q[8] = 0,04$
$q[9] = 0,02$
$q[10] = 0$
$q[11] = 0$
$q[12] = 0$
$q[13] = 0$
$q[14] = 0$
$q[15] = 0$
$q[16] = 0$
$q[17] = 0$
$q[18] = 0,02$
$q[19] = 0,04$
$q[20] = 0$
$q[21] = 0,02$
$q[22] = 0$
$q[23] = 0$
$q[24] = 0$
$q[25] = 0,04$
$q[26] = 0,02$
$q[27] = 0,02$
$q[28] = 0,02$
$q[29] = 0,02$
$q[30] = 0$
$q[31] = 0$
$q[32] = 0,02$

Строка состояния:

$$2.58 < H < 3.26$$

Надлишковість:

```
import math
R = 1 - 2.55 / math.log(33,2)
R1 = 1 - 3.193 / math.log(33,2)
R2 = 1 - 2.04 / math.log(33,2)
R3 = 1 - 2.72 / math.log(33,2)
R4 = 1 - 2.58 / math.log(33,2)
R5 = 1 - 3.26 / math.log(33,2)

print(R,R1,R2,R3,R4,R5)
```

0.49448834891507065 0.36702011689640024 0.5955906791320564 0.46078757217607536 0.4885411530199538 0.35373804606397274

$$0,494 > R(10) > 0,367$$

$$0,595 > R(20) > 0,460$$

$$0,488 > R(30) > 0,353$$

Висновок: виконуючи цю лабораторну роботу ми дізнались про поняття ентропії на символ джерела та його надлишковості.