

МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені Ігоря Сікорського» ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

Криптографія

КОМП’ЮТЕРНИЙ ПРАКТИКУМ

Робота№1

Перевірив:

Виконали:

Чорний О.М.

Студенти групи ФБ-82

Дигас М.В

Кудрик Е.В

Київ – 2020

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.

Завдання для виконання роботи:

1. Уважно прочитати методичні вказівки до виконання комп’ютерного практикуму.
2. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H1 та H2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H1 та H2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H1 та H2 на тому ж тексті, в якому вилучено всі пробіли.
3. За допомогою програми CoolPinkProgram оцінити значення (10) H , (20) H , (30) H .
4. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Виконання роботи:

# Літери з пробілами

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| \_ | 11,608 | о | 8,99 | и | 8,604 |
| е | 7,66 | а | 7,007 | т | 5,776 |
| н | 5,691 | р | 5,198 | с | 4,766 |
| в | 4,281 | м | 2,512 | л | 2,49 |
| п | 2,465 | к | 2,374 | у | 2,336 |
| д | 2,139 | я | 2,013 | з | 1,978 |
| ы | 1,79 | ч | 1,57 | б | 1,365 |
| й | 1,214 | ь | 1,094 | г | 1,02 |
| ц | 0,785 | ш | 0,763 | х | 0,708 |
| ж | 0,655 | ф | 0,581 | щ | 0,236 |
| ю | 0,226 | э | 0,104 |  |  |

# Без пробілів

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| о | 10,171 | и | 9,734 | е | 8,666 |
| а | 7,927 | т | 6,535 | н | 6,438 |
| р | 5,88 | с | 5,392 | в | 4,844 |
| м | 2,842 | л | 2,817 | п | 2,788 |
| к | 2,686 | у | 2,643 | д | 2,42 |
| я | 2,278 | з | 2,238 | ы | 2,025 |
| ч | 1,777 | б | 1,544 | й | 1,373 |
| ь | 1,237 | г | 1,154 | ц | 0,888 |
| ш | 0,864 | х | 0,8 | ж | 0,742 |
| ф | 0,657 | щ | 0,268 | ю | 0,256 |
| э | 0,118 |  |  |  |  |

# Біграми з пробілами, з перетинами

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ст | 2,092 | и\_ | 1,827 | ни | 1,767 | ра | 1,677 | \_п | 1,47 |
| ов | 1,434 | е\_ | 1,412 | \_с | 1,368 | ти | 1,356 | ен | 1,297 |
| ро | 1,256 | \_р | 1,196 | ия | 1,091 | я\_ | 1,086 | об | 1,08 |
| по | 1,047 | \_о | 1,047 | ва | 1,034 | м\_ | 1,034 | то | 0,998 |
| т\_ | 0,987 | аз | 0,958 | ет | 0,921 | ны | 0,915 | но | 0,912 |
| о\_ | 0,847 | \_в | 0,845 | ре | 0,833 | й\_ | 0,821 | на | 0,812 |
| ор | 0,806 | \_и | 0,798 | ан | 0,755 | ие | 0,735 | ль | 0,732 |
| х\_ | 0,708 | ли | 0,681 | ес | 0,68 | ци | 0,664 | в\_ | 0,657 |
| ос | 0,655 | пр | 0,65 | ит | 0,636 | зв | 0,631 | \_н | 0,628 |
| ир | 0,605 | ас | 0,602 | ых | 0,6 | ви | 0,586 | за | 0,583 |
| де | 0,582 | те | 0,557 | а\_ | 0,554 | сп | 0,551 | со | 0,541 |
| \_к | 0,537 | \_у | 0,537 | че | 0,533 | ив | 0,518 | го | 0,51 |
| ко | 0,509 | фо | 0,509 | рм | 0,509 | ь\_ | 0,506 | ии | 0,506 |
| ак | 0,504 | ка | 0,498 | уч | 0,486 | пе | 0,482 | \_ф | 0,48 |
| ац | 0,477 | ки | 0,476 | ом | 0,475 | ал | 0,474 | \_з | 0,471 |
| вы | 0,47 | ел | 0,466 | ае | 0,46 | \_д | 0,455 | же | 0,45 |
| ол | 0,434 | ед | 0,431 | \_м | 0,43 | ру | 0,424 | да | 0,416 |
| ис | 0,416 | ад | 0,407 | во | 0,396 | та | 0,391 | от | 0,39 |
| у\_ | 0,389 | кт | 0,388 | ле | 0,386 | ий | 0,385 | ьн | 0,384 |
| еч | 0,381 | нн | 0,38 | ми | 0,377 | ок | 0,375 | ши | 0,372 |
| од | 0,363 | ей | 0,363 | бр | 0,36 | ше | 0,358 | им | 0,342 |
| он | 0,341 | ча | 0,338 | не | 0,331 | вн | 0,328 | ть | 0,328 |
| \_а | 0,326 | чи | 0,322 | тв | 0,313 | ы\_ | 0,307 | с\_ | 0,306 |
| ят | 0,306 | др | 0,305 | \_ш | 0,302 | ло | 0,296 | мо | 0,294 |
| аж | 0,294 | тр | 0,284 | \_т | 0,284 | кр | 0,283 | ем | 0,281 |
| ач | 0,28 | из | 0,28 | уг | 0,277 | бе | 0,276 | ав | 0,271 |
| бу | 0,269 | ой | 0,263 | ук | 0,261 | зн | 0,261 | ож | 0,241 |
| оз | 0,24 | су | 0,237 | оо | 0,221 | ые | 0,22 | си | 0,21 |
| дл | 0,208 | уе | 0,208 | ог | 0,208 | ме | 0,207 | ма | 0,207 |
| ея | 0,206 | ин | 0,206 | ам | 0,204 | вл | 0,203 | ер | 0,2 |
| дн | 0,197 | яе | 0,193 | ид | 0,192 | пл | 0,187 | ля | 0,187 |
| чт | 0,184 | ту | 0,181 | бо | 0,18 | \_ч | 0,179 | еа | 0,179 |
| сл | 0,175 | сс | 0,174 | ую | 0,174 | зо | 0,173 | жн | 0,173 |
| ве | 0,173 | ву | 0,173 | нс | 0,169 | му | 0,169 | гу | 0,169 |
| ый | 0,169 | ым | 0,168 | га | 0,167 | ев | 0,162 | ю\_ | 0,16 |
| ег | 0,155 | ур | 0,152 | еп | 0,151 | ая | 0,142 | йш | 0,141 |
| гр | 0,141 | ич | 0,14 | чн | 0,139 | ец | 0,137 | иа | 0,137 |
| тс | 0,133 | ул | 0,133 | ьт | 0,133 | ри | 0,131 | ое | 0,131 |
| оп | 0,13 | йн | 0,127 | вс | 0,125 | се | 0,125 | ыс | 0,121 |
| сш | 0,121 | ря | 0,121 | яд | 0,121 | дк | 0,121 | кж | 0,121 |
| оц | 0,12 | це | 0,12 | аб | 0,118 | ры | 0,114 | ла | 0,11 |
| ян | 0,109 | ыв | 0,108 | иг | 0,108 | ну | 0,108 | их | 0,108 |
| бс | 0,105 | дг | 0,105 | вк | 0,105 | ик | 0,105 | ее | 0,105 |
| зл | 0,105 | \_э | 0,104 | ущ | 0,104 | ат | 0,104 | оч | 0,104 |
| ио | 0,102 | оя | 0,101 | аш | 0,101 | еб | 0,1 | нт | 0,1 |
| мы | 0,1 | ди | 0,099 | мк | 0,098 | ар | 0,098 | ищ | 0,098 |
| щи | 0,098 | ус | 0,076 | зы | 0,073 | эт | 0,071 | бл | 0,071 |
| ьк | 0,071 | рг | 0,069 | ап | 0,069 | щн | 0,068 | бн | 0,068 |
| тя | 0,068 | ям | 0,068 | сн | 0,067 | нф | 0,065 | па | 0,065 |
| аг | 0,065 | нд | 0,065 | ск | 0,065 | ып | 0,065 | лн | 0,065 |
| ня | 0,065 | зр | 0,065 | тк | 0,065 | ке | 0,065 | тн | 0,064 |
| йт | 0,059 | яс | 0,055 | \_е | 0,051 | ты | 0,05 | зи | 0,05 |
| иц | 0,05 | ще | 0,037 | фи | 0,037 | дм | 0,037 | ут | 0,037 |
| зд | 0,037 | сф | 0,036 | фе | 0,036 | ду | 0,035 | бы | 0,035 |
| ащ | 0,034 | ща | 0,034 | ай | 0,034 | сь | 0,034 | уд | 0,034 |
| гд | 0,034 | \_г | 0,034 | еш | 0,034 | \_л | 0,034 | лю | 0,034 |
| юб | 0,034 | чу | 0,034 | жи | 0,033 | вш | 0,033 | ша | 0,033 |
| ся | 0,033 | дс | 0,033 | эк | 0,033 | кс | 0,033 | рк | 0,033 |
| ют | 0,033 | \_б | 0,032 | пы | 0,032 | ыт | 0,032 | ош | 0,032 |
| рн | 0,032 | яр | 0,027 | яи | 0,027 | ию | 0,018 | йр | 0,017 |
| йи | 0,013 | йс | 0,013 | яз | 0,013 | яп | 0,011 | ч\_ | 0,009 |
| йп | 0,009 | йз | 0,008 | мт | 0,007 | йе | 0,006 | мс | 0,005 |
| мз | 0,005 | мп | 0,004 | чр | 0,003 | мр | 0,003 | чз | 0,003 |
| чп | 0,003 | чс | 0,003 | уп | 0,002 | мн | 0,002 | уи | 0,002 |
| уз | 0,001 | ун | 0,0 |  |  |  |  |  |  |

# Біграми без пробілів, з перетинами

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ст | 2,367 | | | ни | | | 1,999 | ра | | 1,897 | | ен | | 1,641 | ов | 1,622 | | |
| ти | 1,574 | | | ро | | | 1,421 | то | | 1,306 | | ия | | 1,235 | об | 1,222 | | |
| по | 1,185 | | | ва | | | 1,17 | ес | | 1,111 | | ир | | 1,108 | аз | 1,084 | | |
| ет | 1,081 | | | ор | | | 1,064 | ны | | 1,035 | | но | | 1,032 | ре | 0,942 | | |
| на | 0,919 | | | ан | | | 0,895 | ии | | 0,843 | | ис | | 0,834 | ие | 0,832 | | |
| ль | 0,828 | | | ли | | | 0,77 | ци | | 0,752 | | ос | | 0,741 | пр | 0,736 | | |
| ит | 0,72 | | | ас | | | 0,717 | зв | | 0,714 | | ых | | 0,679 | ви | 0,662 | | |
| за | 0,66 | | | де | | | 0,659 | те | | 0,631 | | сп | | 0,623 | ив | 0,622 | | |
| со | 0,612 | | | че | | | 0,603 | го | | 0,577 | | ко | | 0,576 | фо | 0,575 | | |
| рм | 0,575 | | | им | | | 0,573 | ак | | 0,57 | | мо | | 0,565 | ка | 0,564 | | |
| уч | 0,549 | | | тв | | | 0,549 | пе | | 0,545 | | ац | | 0,54 | ки | 0,539 | | |
| ом | | 0,537 | | ал | 0,537 | | | вы | | | 0,532 | ел | 0,527 | | ед | | | 0,526 |
| во | | 0,521 | | ае | 0,52 | | | же | | | 0,509 | ер | 0,5 | | од | | | 0,491 |
| ол | | 0,49 | | ук | 0,487 | | | ру | | | 0,479 | ми | 0,474 | | да | | | 0,471 |
| ок | | 0,464 | | ад | 0,461 | | | та | | | 0,443 | оо | 0,442 | | от | | | 0,441 |
| кт | | 0,439 | | ле | 0,437 | | | ий | | | 0,436 | ьн | 0,434 | | еч | | | 0,431 |
| нн | | 0,43 | | ши | 0,421 | | | ем | | | 0,417 | ей | 0,41 | | бр | | | 0,408 |
| ше | | 0,406 | | ят | 0,404 | | | ач | | | 0,398 | ин | 0,387 | | он | | | 0,386 |
| ев | | 0,385 | | ма | 0,384 | | | ча | | | 0,382 | оп | 0,378 | | не | | | 0,375 |
| вн | | 0,371 | | ть | 0,371 | | | чи | | | 0,366 | хп | 0,361 | | ву | | | 0,35 |
| др | | 0,345 | | ло | 0,335 | | | аж | | | 0,332 | вф | 0,323 | | тр | | | 0,321 |
| кр | | 0,321 | | из | 0,316 | | | уг | | | 0,313 | бе | 0,312 | | оз | | | 0,311 |
| ав | | 0,307 | | бу | 0,305 | | | ой | | | 0,297 | зн | 0,295 | | сш | | | 0,287 |
| яс | | 0,285 | | ож | 0,273 | | | су | | | 0,268 | ид | 0,258 | | вс | | | 0,256 |
| ат | | 0,254 | | ту | 0,251 | | | ус | | | 0,25 | ые | 0,249 | | ип | | | 0,243 |
| си | | 0,238 | | уе | 0,237 | | | ме | | | 0,236 | дл | 0,236 | | ог | | | 0,235 |
| ио | | 0,233 | | ея | 0,233 | | | ам | | | 0,231 | яе | 0,23 | | вл | | | 0,229 |
| тс | | 0,229 | | дн | 0,223 | | | ян | | | 0,215 | пл | 0,212 | | чт | | | 0,211 |
| ля | | 0,211 | | еп | 0,209 | | | бо | | | 0,204 | еи | 0,203 | | еа | | | 0,202 |
| сл | | 0,198 | | сс | 0,197 | | | ую | | | 0,197 | йс | 0,197 | | иа | | | 0,196 |
| зо | | 0,196 | | жн | 0,196 | | | ьв | | | 0,196 | ве | 0,195 | | яп | | | 0,194 |
| нс | | 0,191 | | тш | 0,191 | | | му | | | 0,191 | гу | 0,191 | | ый | | | 0,191 |
| ым | | 0,19 | | га | 0,189 | | | як | | | 0,188 | ур | 0,181 | | се | | | 0,18 |
| яв | | 0,175 | | ег | 0,175 | | | оу | | | 0,165 | ыо | 0,164 | | мп | | | 0,163 |
| ая | | 0,161 | | ап | 0,161 | | | хф | | | 0,16 | йш | 0,16 | | гр | | | 0,159 |
| ич | | 0,159 | | чн | 0,158 | | | тп | | | 0,156 | ик | 0,155 | | ец | | | 0,155 |
| ям | | 0,154 | | ул | 0,15 | | | ьт | | | 0,15 | ар | 0,15 | | йн | | | 0,15 |
| тн | | 0,149 | | ри | 0,148 | | | мк | | | 0,148 | ое | 0,148 | | йр | | | 0,147 |
| яи | | 0,145 | | йп | 0,137 | | | ыс | | | 0,137 | ря | 0,137 | | яд | | | 0,137 |
| дк | | 0,137 | | аа | 0,137 | | | кж | | | 0,137 | оц | 0,136 | | це | | | 0,136 |
| аб | | 0,134 | | ып | 0,13 | | | ры | | | 0,129 | мр | 0,126 | | мн | | | 0,126 |
| ла | | 0,124 | | мм | 0,124 | | | ыв | | | 0,122 | иг | 0,122 | | ну | | | 0,122 |
| юр | | 0,122 | | их | 0,122 | | | сд | | | 0,121 | хз | 0,119 | | ьк | | | 0,119 |
| бс | | 0,119 | | дг | 0,119 | | | вк | | | 0,119 | ее | 0,119 | | зл | | | 0,119 |
| мд | | 0,119 | | ущ | 0,118 | | | оч | | | 0,118 | йд | 0,115 | | оя | | | 0,114 |
| аш | | 0,114 | | еб | 0,113 | | | нт | | | 0,113 | мы | 0,113 | | аг | | | 0,112 |
| ди | | 0,112 | | ищ | 0,111 | | | щи | | | 0,111 | тк | 0,111 | | йи | | | 0,11 |
| ьз | | 0,1 | | йт | 0,093 | | | зы | | | 0,082 | тч | 0,082 | | ьэ | | | 0,081 |
| эт | | 0,081 | | бл | 0,081 | | | рг | | | 0,078 | ху | 0,078 | | щн | | | 0,076 |
| бн | | 0,076 | | тя | 0,076 | | | сн | | | 0,076 | яр | 0,076 | | иу | | | 0,076 |
| нф | | 0,074 | | па | 0,074 | | | нд | | | 0,074 | ск | 0,074 | | ео | | | 0,074 |
| лн | | 0,074 | | ня | 0,074 | | | ез | | | 0,074 | зр | 0,074 | | ке | | | 0,074 |
| йз | | 0,057 | | ты | 0,057 | | | зи | | | 0,057 | иц | 0,057 | | ут | | | 0,055 |
| мт | | | 0,053 | ыу | | 0,047 | | еу | 0,042 | | | ыи | 0,042 | | ще | | 0,042 | |
| фи | | | 0,042 | хи | | 0,042 | | дм | 0,042 | | | йв | 0,041 | | вз | | 0,041 | |
| зд | | | 0,041 | ьо | | 0,041 | | хн | 0,041 | | | йк | 0,04 | | сф | | 0,04 | |
| фе | | | 0,04 | йа | | 0,04 | | ду | 0,04 | | | бы | 0,04 | | ьч | | 0,04 | |
| иф | | | 0,04 | ыр | | 0,038 | | ув | 0,038 | | | ои | 0,038 | | яо | | 0,038 | |
| ащ | | | 0,038 | ща | | 0,038 | | ай | 0,038 | | | сь | 0,038 | | ьс | | 0,038 | |
| уд | | | 0,038 | гд | | 0,038 | | еш | 0,038 | | | ьл | 0,038 | | лю | | 0,038 | |
| юб | | | 0,038 | юз | | 0,038 | | чу | 0,038 | | | яз | 0,038 | | жи | | 0,037 | |
| вш | | | 0,037 | ша | | 0,037 | | ся | 0,037 | | | ао | 0,037 | | дс | | 0,037 | |
| йэ | | | 0,037 | эк | | 0,037 | | кс | 0,037 | | | рк | 0,037 | | ют | | 0,037 | |
| иб | | | 0,037 | йо | | 0,037 | | пы | 0,037 | | | ыт | 0,037 | | ош | | 0,036 | |
| вв | | | 0,036 | тз | | 0,036 | | св | 0,036 | | | рн | 0,036 | | оф | | 0,02 | |
| ию | | | 0,02 | юп | | 0,02 | | йе | 0,012 | | | мс | 0,011 | | уи | | 0,01 | |
| мз | | | 0,009 | уп | | 0,008 | | уз | 0,008 | | | чр | 0,006 | | чз | | 0,005 | |
| чп | | | 0,005 | чс | | 0,004 | | ун | 0,003 | | |  |  | |  | |  | |

# Біграми без пробілів, з перетинами

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ст | 2,432 | | ни | | 1,988 | ра | | 1,882 | ов | | 1,644 | ен | 1,616 | |
| ти | 1,542 | | ро | | 1,394 | то | | 1,301 | ия | | 1,232 | по | 1,205 | |
| об | 1,187 | | ва | | 1,128 | ир | | 1,127 | ес | | 1,071 | аз | 1,064 | |
| ет | 1,058 | | ор | | 1,05 | но | | 1,05 | ны | | 1,04 | на | 0,936 | |
| ре | 0,927 | | ан | | 0,905 | ии | | 0,896 | ль | | 0,844 | ие | 0,825 | |
| ис | 0,824 | | ли | | 0,766 | пр | | 0,761 | ци | | 0,747 | ит | 0,731 | |
| зв | 0,72 | | ас | | 0,712 | ос | | 0,697 | ых | | 0,671 | за | 0,649 | |
| ви | 0,647 | | те | | 0,642 | де | | 0,638 | ив | | 0,637 | сп | 0,63 | |
| че | 0,624 | | со | | 0,599 | фо | | 0,598 | рм | | 0,598 | ак | 0,59 | |
| го | 0,582 | | им | | 0,573 | ка | | 0,571 | ко | | 0,569 | мо | 0,566 | |
| тв | 0,563 | | ом | | 0,548 | ац | | 0,541 | уч | | 0,541 | ед | 0,54 | |
| ае | 0,538 | | вы | | 0,534 | пе | | 0,531 | во | | 0,525 | ал | 0,519 | |
| ел | 0,516 | | ки | | 0,513 | же | | 0,513 | ру | | 0,505 | од | 0,496 | |
| ол | 0,495 | | да | | 0,486 | ук | | 0,479 | ер | | 0,478 | ад | 0,47 | |
| ок | 0,465 | | оо | | 0,458 | ле | | 0,452 | кт | | 0,448 | ми | 0,446 | |
| нн | 0,443 | | еч | | 0,441 | ий | | 0,434 | ем | | 0,427 | та | 0,426 | |
| бр | 0,417 | | от | | 0,413 | ьн | | 0,41 | ши | | 0,41 | не | 0,399 | |
| ей | 0,399 | | ше | | 0,396 | ят | | 0,394 | ач | | 0,389 | ев | 0,387 | |
| ин | 0,386 | | ма | | 0,386 | он | | 0,384 | оп | | 0,382 | ча | 0,382 | |
| ть | 0,365 | | вн | | 0,363 | хп | | 0,362 | чи | | 0,359 | аж | 0,349 | |
| ло | 0,339 | | ву | | 0,336 | др | | 0,33 | тр | | 0,323 | ав | 0,322 | |
| бе | 0,321 | | кр | | 0,319 | из | | 0,317 | бу | | 0,316 | оз | 0,314 | |
| вф | 0,301 | | уг | | 0,297 | сш | | 0,293 | ой | | 0,29 | зн | 0,289 | |
| яс | 0,279 | | су | | 0,276 | вс | | 0,275 | ат | | 0,271 | ож | 0,267 | |
| ид | 0,261 | | ые | | 0,248 | си | | 0,247 | уе | | 0,246 | ип | 0,246 | |
| ея | 0,244 | | ме | | 0,241 | ог | | 0,241 | ус | | 0,238 | ту | 0,236 | |
| дл | 0,231 | | ио | | 0,229 | яе | | 0,227 | тс | | 0,226 | ам | 0,224 | |
| вл | 0,218 | | ля | | 0,216 | дн | | 0,216 | чт | | 0,215 | еп | 0,212 | |
| еи | | 0,211 | иа | 0,21 | | ян | 0,209 | | сл | 0,207 | | бо | | 0,207 |
| ьв | | 0,205 | ый | 0,205 | | йс | 0,202 | | тш | 0,202 | | гу | | 0,202 |
| ве | | 0,201 | пл | 0,201 | | як | 0,199 | | сс | 0,198 | | еа | | 0,196 |
| зо | | 0,196 | нс | 0,196 | | ур | 0,19 | | га | 0,19 | | яп | | 0,188 |
| ым | | 0,188 | ег | 0,186 | | жн | 0,186 | | ую | 0,184 | | му | | 0,181 |
| оу | | 0,175 | гр | 0,173 | | ая | 0,172 | | йш | 0,171 | | ыо | | 0,171 |
| се | | 0,17 | тп | 0,164 | | яв | 0,164 | | ец | 0,163 | | хф | | 0,162 |
| ич | | 0,159 | мк | 0,159 | | чн | 0,155 | | ям | 0,151 | | мп | | 0,151 |
| ар | | 0,15 | ул | 0,15 | | ьт | 0,15 | | ое | 0,148 | | ап | | 0,147 |
| ик | | 0,147 | ри | 0,146 | | тн | 0,146 | | оц | 0,145 | | яд | | 0,144 |
| йр | | 0,143 | йн | 0,142 | | яи | 0,136 | | ну | 0,134 | | юр | | 0,134 |
| йп | | 0,132 | аб | 0,132 | | мн | 0,13 | | ыс | 0,129 | | ря | | 0,129 |
| дк | | 0,129 | аа | 0,129 | | кж | 0,129 | | ла | 0,129 | | бс | | 0,128 |
| вк | | 0,128 | их | 0,128 | | ып | 0,128 | | сд | 0,128 | | це | | 0,127 |
| ее | | 0,125 | зл | 0,125 | | мд | 0,125 | | мр | 0,123 | | ьк | | 0,122 |
| ры | | 0,121 | оя | 0,121 | | хз | 0,12 | | еб | 0,12 | | мм | | 0,119 |
| йд | | 0,119 | ыв | 0,119 | | оч | 0,116 | | ищ | 0,115 | | нт | | 0,112 |
| мы | | 0,112 | тк | 0,112 | | йи | 0,111 | | ущ | 0,11 | | дг | | 0,11 |
| иг | | 0,109 | аг | 0,108 | | ди | 0,108 | | аш | 0,108 | | щи | | 0,107 |
| ьз | | 0,106 | зы | 0,089 | | йт | 0,089 | | ху | 0,085 | | иу | | 0,082 |
| эт | | 0,082 | яр | 0,081 | | ьэ | 0,08 | | бл | 0,08 | | щн | | 0,08 |
| сн | | 0,078 | ня | 0,077 | | рг | 0,076 | | тч | 0,076 | | па | | 0,074 |
| нд | | 0,074 | ео | 0,074 | | нф | 0,073 | | ск | 0,073 | | бн | | 0,073 |
| тя | | 0,073 | лн | 0,071 | | ез | 0,071 | | зр | 0,071 | | ке | | 0,071 |
| зи | | 0,062 | йз | 0,059 | | ты | 0,058 | | ут | 0,057 | | мт | | 0,052 |
| иц | | 0,051 | фе | 0,046 | | йа | 0,046 | | еу | 0,046 | | ыи | | 0,046 |
| ще | | 0,046 | фи | 0,046 | | дм | 0,046 | | ыу | 0,045 | | ьч | | 0,043 |
| вз | | 0,043 | яо | 0,042 | | ащ | 0,042 | | ай | 0,042 | | сь | | 0,042 |
| уд | | 0,042 | лю | 0,042 | | юз | 0,042 | | йв | 0,04 | | зд | | 0,04 |
| жи | | 0,039 | вш | 0,039 | | ся | 0,039 | | ао | 0,039 | | йэ | | 0,039 |
| кс | | 0,039 | вв | 0,039 | | тз | 0,039 | | св | 0,039 | | ош | | 0,039 |
| яз | | 0,038 | иб | 0,038 | | ыт | 0,038 | | ьо | 0,037 | | хи | | 0,037 |
| ыр | | 0,036 | ду | 0,036 | | бы | 0,036 | | иф | 0,036 | | хн | | 0,036 |
| ша | | 0,036 | дс | 0,036 | | эк | 0,036 | | рк | 0,036 | | йо | | 0,035 |
| пы | | 0,035 | ув | 0,035 | | ои | 0,035 | | ща | 0,035 | | ьс | | 0,035 |
| гд | | 0,035 | еш | 0,035 | | ьл | 0,035 | | юб | 0,035 | | чу | | 0,035 |
| йк | | 0,034 | сф | 0,034 | | ют | 0,034 | | рн | 0,034 | | ию | | 0,023 |
| оф | | 0,017 | юп | 0,017 | | йе | 0,014 | | мс | 0,011 | | уп | | 0,008 |
| мз | | 0,008 | уз | 0,008 | | чр | 0,007 | | уи | 0,007 | | чз | | 0,006 |
| чп | | 0,004 | чс | 0,003 | | ун | 0,002 | |  |  | |  | |  |

# Біграми з пробілами, з перетинами

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ст | 2,103 | | и\_ | 1,827 | | ни | | 1,754 | ра | | 1,684 | \_п | 1,458 | |
| ов | 1,427 | | е\_ | 1,411 | | ти | | 1,362 | \_с | | 1,354 | ен | 1,305 | |
| ро | 1,266 | | \_р | 1,192 | | ия | | 1,092 | я\_ | | 1,089 | об | 1,081 | |
| по | 1,075 | | \_о | 1,04 | | ва | | 1,035 | м\_ | | 1,025 | т\_ | 1,004 | |
| то | 0,996 | | аз | 0,96 | | ет | | 0,907 | но | | 0,896 | ны | 0,888 | |
| о\_ | 0,853 | | ре | 0,84 | | \_в | | 0,833 | й\_ | | 0,816 | ор | 0,813 | |
| \_и | 0,801 | | на | 0,801 | | ан | | 0,762 | ие | | 0,741 | ль | 0,721 | |
| х\_ | 0,694 | | ци | 0,68 | | ли | | 0,677 | ес | | 0,674 | в\_ | 0,673 | |
| ос | 0,655 | | ит | 0,643 | | зв | | 0,642 | пр | | 0,638 | \_н | 0,636 | |
| ых | 0,618 | | ас | 0,605 | | ир | | 0,595 | за | | 0,584 | де | 0,583 | |
| ви | 0,576 | | а\_ | 0,564 | | со | | 0,556 | \_к | | 0,552 | сп | 0,547 | |
| те | 0,544 | | \_у | 0,527 | | че | | 0,523 | ив | | 0,519 | ко | 0,511 | |
| го | 0,51 | | ь\_ | 0,506 | | фо | | 0,505 | рм | | 0,505 | ак | 0,5 | |
| ии | 0,496 | | уч | 0,492 | | ал | | 0,489 | \_ф | | 0,488 | ка | 0,487 | |
| пе | 0,483 | | ки | 0,481 | | вы | | 0,475 | \_з | | 0,472 | ом | 0,47 | |
| ел | 0,469 | | ац | 0,469 | | ае | | 0,463 | \_д | | 0,463 | ру | 0,447 | |
| же | 0,441 | | ол | 0,439 | | \_м | | 0,438 | ед | | 0,422 | ад | 0,421 | |
| ис | 0,414 | | от | 0,407 | | да | | 0,403 | ий | | 0,401 | кт | 0,398 | |
| ьн | 0,396 | | ле | 0,393 | | во | | 0,392 | та | | 0,391 | нн | 0,387 | |
| ши | 0,385 | | еч | 0,385 | | ми | | 0,378 | ей | | 0,377 | у\_ | 0,373 | |
| ок | 0,362 | | од | 0,361 | | бр | | 0,36 | ше | | 0,36 | он | 0,35 | |
| им | 0,349 | | вн | 0,34 | | не | | 0,336 | ча | | 0,331 | ть | 0,325 | |
| ы\_ | 0,324 | | \_а | 0,316 | | чи | | 0,315 | ят | | 0,314 | тв | 0,313 | |
| с\_ | 0,3 | | др | 0,297 | | ач | | 0,293 | \_ш | | 0,292 | мо | 0,289 | |
| аж | 0,288 | | ем | 0,282 | | бе | | 0,281 | ло | | 0,278 | \_т | 0,278 | |
| тр | 0,273 | | из | 0,272 | | ав | | 0,272 | кр | | 0,269 | зн | 0,268 | |
| бу | 0,262 | | ой | 0,261 | | уг | | 0,261 | ук | | 0,259 | ож | 0,252 | |
| су | 0,245 | | оз | 0,232 | | ые | | 0,219 | си | | 0,219 | дл | 0,217 | |
| оо | | 0,212 | уе | | 0,211 | ам | 0,206 | | ер | 0,206 | | ме | | 0,205 |
| ог | | 0,204 | ин | | 0,204 | ея | 0,201 | | яе | 0,2 | | ма | | 0,2 |
| ид | | 0,199 | вл | | 0,196 | дн | 0,193 | | пл | 0,193 | | \_ч | | 0,185 |
| еа | | 0,182 | ля | | 0,18 | сл | 0,18 | | му | 0,178 | | гу | | 0,178 |
| чт | | 0,178 | ве | | 0,176 | ую | 0,176 | | ым | 0,174 | | жн | | 0,174 |
| ту | | 0,173 | бо | | 0,171 | сс | 0,17 | | га | 0,169 | | зо | | 0,168 |
| ый | | 0,168 | ву | | 0,167 | ю\_ | 0,164 | | нс | 0,164 | | ег | | 0,16 |
| ев | | 0,16 | ур | | 0,159 | еп | 0,145 | | ая | 0,142 | | гр | | 0,14 |
| чн | | 0,139 | ич | | 0,137 | тс | 0,135 | | оп | 0,134 | | йш | | 0,134 |
| ул | | 0,131 | ьт | | 0,131 | ое | 0,131 | | ри | 0,13 | | вс | | 0,129 |
| ец | | 0,128 | иа | | 0,128 | сш | 0,126 | | ря | 0,126 | | дк | | 0,126 |
| кж | | 0,126 | оц | | 0,124 | аб | 0,122 | | се | 0,12 | | йн | | 0,12 |
| це | | 0,117 | ыс | | 0,116 | яд | 0,116 | | ла | 0,11 | | иг | | 0,11 |
| ну | | 0,11 | оч | | 0,11 | ик | 0,11 | | ры | 0,108 | | бс | | 0,108 |
| дг | | 0,108 | ян | | 0,108 | \_э | 0,104 | | ыв | 0,104 | | их | | 0,103 |
| вк | | 0,103 | аш | | 0,103 | оя | 0,102 | | еб | 0,102 | | ио | | 0,101 |
| мы | | 0,1 | ди | | 0,1 | щи | 0,1 | | ее | 0,1 | | зл | | 0,1 |
| нт | | 0,099 | мк | | 0,098 | ат | 0,098 | | ущ | 0,097 | | ар | | 0,097 |
| ищ | | 0,097 | зы | | 0,075 | ьк | 0,073 | | ус | 0,072 | | щн | | 0,07 |
| бн | | 0,07 | тя | | 0,07 | эт | 0,069 | | бл | 0,069 | | ап | | 0,069 |
| лн | | 0,069 | зр | | 0,069 | ке | 0,069 | | рг | 0,066 | | аг | | 0,065 |
| ск | | 0,065 | сн | | 0,065 | нф | 0,065 | | па | 0,065 | | нд | | 0,065 |
| ям | | 0,065 | тн | | 0,064 | ып | 0,062 | | ня | 0,062 | | тк | | 0,062 |
| яс | | 0,057 | йт | | 0,055 | \_е | 0,051 | | иц | 0,051 | | зи | | 0,049 |
| ты | | 0,048 | ще | | 0,041 | ут | 0,041 | | ду | 0,037 | | бы | | 0,037 |
| фе | | 0,037 | зд | | 0,037 | ша | 0,035 | | эк | 0,035 | | ащ | | 0,034 |
| ай | | 0,034 | сь | | 0,034 | \_г | 0,034 | | \_л | 0,034 | | юб | | 0,034 |
| сф | | 0,034 | ща | | 0,034 | уд | 0,034 | | гд | 0,034 | | еш | | 0,034 |
| лю | | 0,034 | чу | | 0,034 | \_б | 0,033 | | пы | 0,033 | | фи | | 0,033 |
| дм | | 0,033 | ыт | | 0,032 | жи | 0,031 | | вш | 0,031 | | ся | | 0,031 |
| дс | | 0,031 | кс | | 0,031 | рк | 0,031 | | рн | 0,031 | | яр | | 0,03 |
| ош | | 0,029 | ют | | 0,029 | яи | 0,022 | | йс | 0,016 | | ию | | 0,016 |
| йр | | 0,015 | яз | | 0,012 | йи | 0,012 | | яп | 0,01 | | мт | | 0,008 |
| ч\_ | | 0,008 | йз | | 0,007 | йп | 0,007 | | йе | 0,006 | | мс | | 0,005 |
| мз | | 0,005 | мп | | 0,004 | мр | 0,004 | | чр | 0,004 | | чз | | 0,003 |
| уп | | 0,003 | чс | | 0,003 | чп | 0,003 | | уи | 0,002 | | мн | | 0,002 |
| уз | | 0,001 |  | |  |  |  | |  |  | |  | |  |

# Таблиці зі значеннями H1, H2

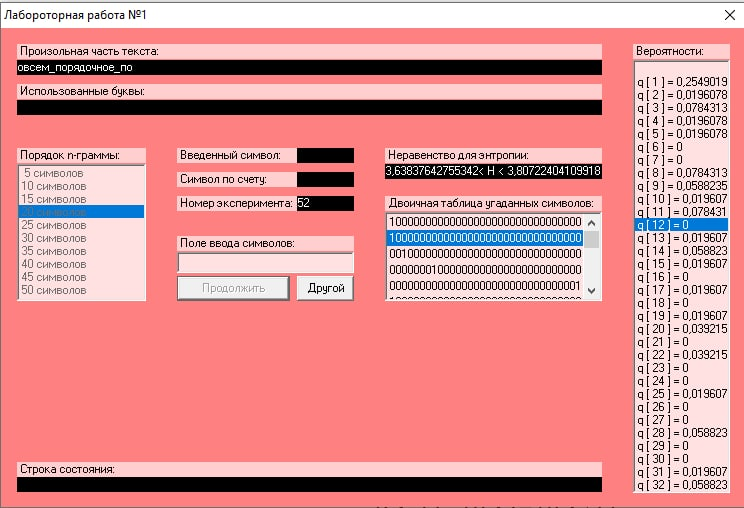
# H1(літери)

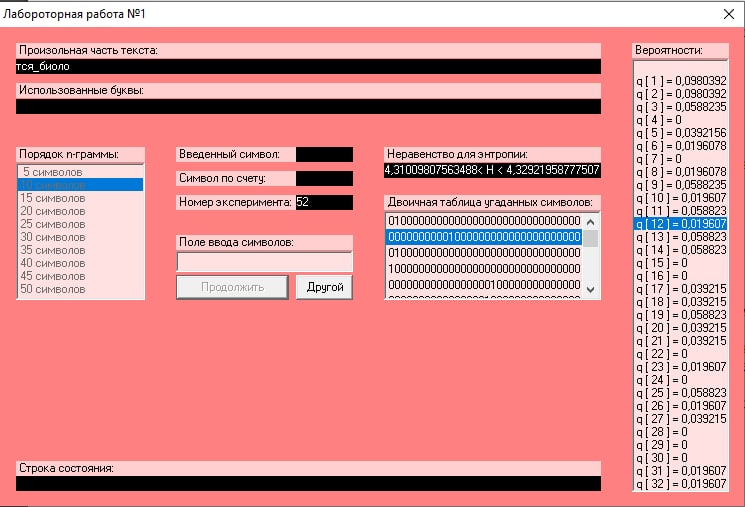
|  |  |
| --- | --- |
| З пробілами | Без пробілів |
| 4.424591 | 4.419631 |

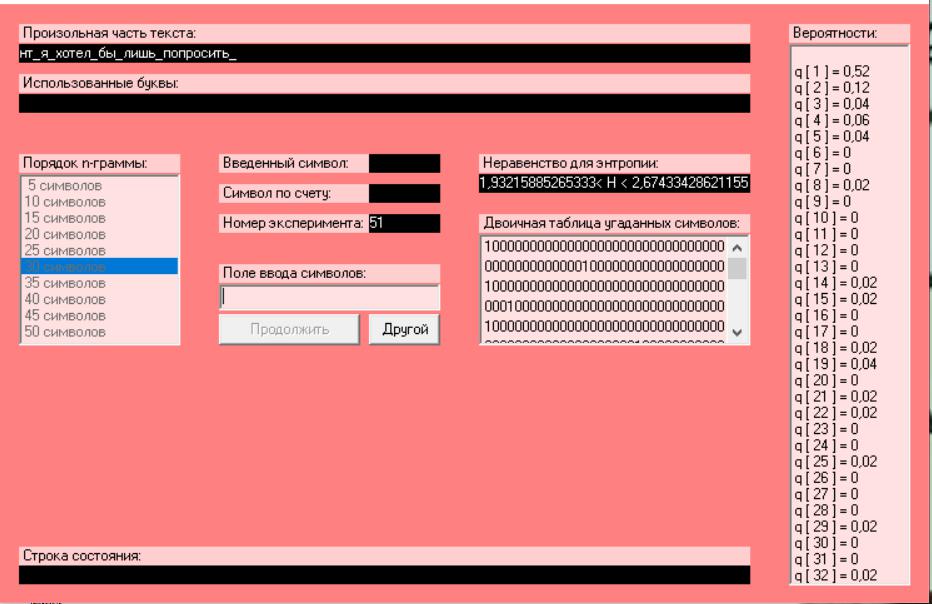
# H2 (Біграми)

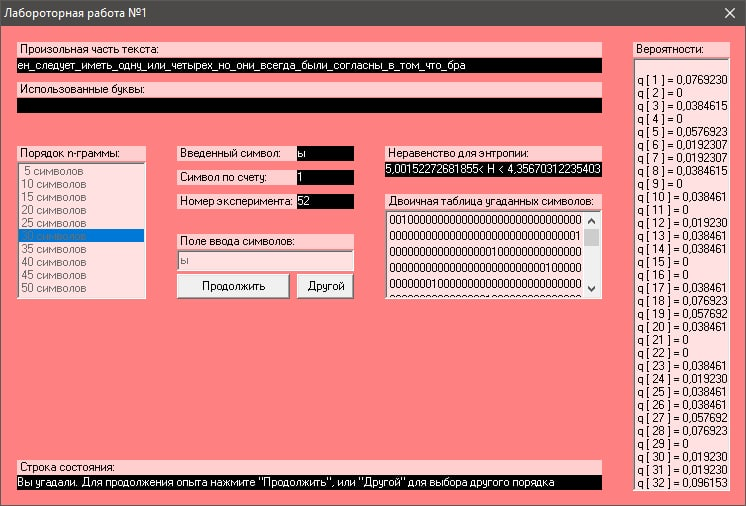
|  |  |  |
| --- | --- | --- |
|  | З пробілами | Без пробілів |
| З перетинами | 3.792291 | 3.832963 |
| Без перетинів | 3.793215 | 3.83233 |

# За допомогою програми CoolPinkProgram оцінюю значення H(10), H(20), H(30) .









Для зручності створимо таблицю

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| H(10) | | H(20) | | H(30) | |
| Нижня межа | Верхня межа | Нижня межа | Верхня межа | Нижня межа | Верхня межа |
| H(10)1 | H(10)2 | H(20)1 | H(20)2 | H(30)1 | H(30)2 |
| 4.31009807563488 | 4.32921958777507 | 3.63837642755342 | 3.80722404109918 | 5.00152272681855 | 4.35670312235403 |

Висновок: Під час використання программи CoolPinkProgram зрозуміли, що чим більше символів у певній послідовності нам відомо, тим з більшою ймовірністю ми вгадаємо наступний символ, оскільки значення ентропіїї зменшується, при збільшенні відомих символів.