

Міністерство освіти і науки України  
Національний технічний університет України  
"Київський політехнічний інститут імені Ігоря Сікорського"  
Фізико-технічний інститут

**Лабораторна робота №1**  
ЕКСПЕРИМЕНТАЛЬНА ОЦІНКА ЕНТРОПІЇ НА СИМВОЛ ДЖЕРЕЛА  
ВІДКРИТОГО ТЕКСТУ

Виконали: студенти гр. ФІ-94,  
Молдован Д.В., Пелешенко Л.І.

Перевірив: Чорний О.М.

Київ - 2022

### **Мета лабораторної 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.

### **Постановка задачі:**

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.

### **Опис труднощів:**

- Виникали проблеми з формулою для обрахунку ентропії: складно було розібратись та спочатку вона була реалізована з помилками.
- Функції, які обраховують частоти для біграм з перетином і без перетину майже ідентичні (відрізняється лише рядок з циклом). Тому в кодї є фрагмент, який двічі повторюється (за винятком рядку з циклом). Існують способи вирішення такої проблеми, адже дана реалізація робить код більш громіздким, але він все рівно є робочим.
- Пошук найкращих біграм робився просто сортуванням, реалізація якого була знайдена в мережі інтернет.

### **Результат роботи програми:**

Ентропія тексту з пробілами;

Ентропія тексту з пробілами та перетином;

Ентропія тексту з пробілами та без перетину;

Ентропія тексту без пробілів;

Ентропія тексту без пробілів та перетину;

Ентропія тексту без пробілів та без перетину відповідно.

```
H1 with gaps = 4.447975437586149
H2 with gaps with intersection = 3.974999153773363
H2 with gaps without intersection = 3.933141321045024
H1 no gaps = 4.421720836108205
H2 no gaps with intersection = 4.020811911404876
H2 no gaps without intersection = 3.974883235762987
```

Біграми, які трапляються найчастіше:

```
Best bigrams [
  [ 'ст', 0.02063524175601861 ],
  [ 'но', 0.017600647380133522 ],
  [ 'ра', 0.013554521545620068 ],
  [ 'на', 0.011936071211814687 ],
  [ 'то', 0.011733764920089014 ],
  [ 'ов', 0.011531458628363342 ]
]
```

Алфавіт без пробіла:

Частота літер:

```
Table for literals {
  'а': 0.07605177993527508,
  'г': 0.014563106796116505,
  'о': 0.1063915857605178,
  'д': 0.030137540453074433,
  'ы': 0.019417475728155338,
  'ш': 0.0056634304207119745,
  'л': 0.040453074433656956,
  'и': 0.08272653721682847,
  'б': 0.011326860841423949,
  'с': 0.061893203883495146,
  'т': 0.06897249190938512,
  'р': 0.0505663430420712,
  'н': 0.07423139158576052,
  'е': 0.08980582524271845,
  'к': 0.028923948220064725,
  'п': 0.022451456310679612,
  'ж': 0.007079288025889968,
  'в': 0.041868932038834954,
  'м': 0.030946601941747573,
  'ь': 0.01820388349514563,
  'э': 0.017192556634304208,
  'й': 0.012540453074433657,
  'у': 0.02063106796116505,
  'х': 0.011124595469255663,
  'ц': 0.007281553398058253,
  'щ': 0.0034385113268608413,
  'я': 0.01779935275080906,
  'ч': 0.013349514563106795,
  'ю': 0.007483818770226537,
  'э': 0.00465210355987055,
  'ф': 0.0028317152103559872
}
```

## Частота біграм з перетином:

Table for bigrams {	'нв': 0.00829455796075258,	'вн': 0.0028322880841594174,	'цп': 0.00020230629172567267,
'ар': 0.000606918875177018,	'ые': 0.00303459437588509,	'еш': 0.00040461258345134534,	'ещ': 0.000606918875177018,
'го': 0.005866882460044507,	'ев': 0.00303459437588509,	'ем': 0.005259963584867489,	'ша': 0.00020230629172567267,
'од': 0.009306089419380943,	'во': 0.005664576168318835,	'мв': 0.000606918875177018,	'ав': 0.004653044709690472,
'ды': 0.0010115314586283633,	'ып': 0.00020230629172567267,	'ов': 0.011531458628363342,	'вш': 0.0010115314586283633,
'ьш': 0.000606918875177018,	'пр': 0.00687841391867287,	'ну': 0.000606918875177018,	'ши': 0.001213837750354036,
'шл': 0.0008092251669026907,	'от': 0.009508395711106616,	'ут': 0.0014161440420797087,	'их': 0.00303459437588509,
'ли': 0.007283026502124216,	'те': 0.00687841391867287,	'ре': 0.006271495043495853,	'ом': 0.008092251669026907,
'ид': 0.002427675500708072,	'ек': 0.0028322880841594174,	'нн': 0.005259963584867489,	'мс': 0.001213837750354036,
'да': 0.005259963584867489,	'ал': 0.00829455796075258,	'йб': 0.000606918875177018,	'та': 0.0038438195427877805,
'аш': 0.000606918875177018,	'ла': 0.007687639085575561,	'бо': 0.0028322880841594174,	'ах': 0.0016184503338053813,
'иб': 0.000606918875177018,	'ам': 0.0020230629172567267,	'ор': 0.006473801335221525,	'хо': 0.004248432126239126,
'бы': 0.0014161440420797087,	'мо': 0.004046125834513453,	'рь': 0.00020230629172567267,	'вы': 0.0016184503338053813,
'ыс': 0.0014161440420797087,	'ол': 0.005664576168318835,	'ьб': 0.00040461258345134534,	'ых': 0.00020230629172567267,
'ст': 0.02063524175601861,	'ло': 0.004450738417964798,	'ей': 0.001820756625531054,	'он': 0.008901476835929597,
'тр': 0.009306089419380943,	'до': 0.0032369006676107627,	'ит': 0.007080720210398544,	'на': 0.011936071211814687,
'ро': 0.010317620878009305,	'ос': 0.009306089419380943,	'ог': 0.005866882460044507,	'со': 0.008901476835929597,
'ои': 0.002225369208982399,	'ть': 0.00606918875177018,	'ге': 0.00040461258345134534,	'ан': 0.0038438195427877805,
'ин': 0.005664576168318835,	'ье': 0.0016184503338053813,	'еп': 0.0028322880841594174,	'ош': 0.0008092251669026907,
'не': 0.010722233461460652,	'ел': 0.005866882460044507,	'др': 0.00040461258345134534,	'ас': 0.009508395711106616,
'ес': 0.010519927169734978,	'л': 0.009508395711106616,	'ру': 0.00303459437588509,	'сь': 0.004450738417964798,
'сл': 0.004653044709690472,	'ен': 0.014768359295974105,	'уг': 0.0010115314586283633,	'ьн': 0.004248432126239126,
'лы': 0.00040461258345134534,	'ыв': 0.001213837750354036,	'гу': 0.0010115314586283633,	'ни': 0.011126846044911997,
'шн': 0.000606918875177018,	'ьб': 0.00040461258345134534,	'ун': 0.001213837750354036,	'ис': 0.007283026502124216,
'но': 0.017600647380133522,	'бе': 0.002427675500708072,	'ее': 0.004653044709690472,	'дн': 0.004046125834513453,
'ок': 0.0020230629172567267,	'ез': 0.0020230629172567267,	'еб': 0.002225369208982399,	'ой': 0.004855351001416144,
'ка': 0.005462269876593162,	'эд': 0.0010115314586283633,	'ыл': 0.0010115314586283633,	'йр': 0.0010115314586283633,
'ак': 0.004248432126239126,	'де': 0.005462269876593162,	'из': 0.005866882460044507,	'ди': 0.003641513251062108,
'кп': 0.00040461258345134534,	'ей': 0.00303459437588509,	'зо': 0.0008092251669026907,	'ль': 0.004855351001416144,
'по': 0.00910378312765527,	'йс': 0.0028322880841594174,	'вс': 0.0028322880841594174,	'ьс': 0.002427675500708072,
'дс': 0.001213837750354036,	'тв': 0.004248432126239126,	'се': 0.0032369006676107627,	'ск': 0.005664576168318835,
'сн': 0.001820756625531054,	'ви': 0.0050576572931418165,	'ех': 0.001213837750354036,	'ая': 0.000606918875177018,
'еж': 0.0014161440420797087,	'ии': 0.00606918875177018,	'хд': 0.0010115314586283633,	'яв': 0.002225369208982399,
'жн': 0.0014161440420797087,	'ив': 0.005259963584867489,	'иц': 0.0008092251669026907,	'вл': 0.001820756625531054,
'ны': 0.00829455796075258,	'вн': 0.0028322880841594174,	'цп': 0.00020230629172567267,	'ик': 0.004450738417964798,
	'еш': 0.00040461258345134534,	'ещ': 0.000606918875177018,	'ко': 0.009508395711106616,

'ко': 0.009508395711106616,	'св': 0.001820756625531054,	'то': 0.011733764920089014,	'йо': 0.000606918875177018,
'гд': 0.0008092251669026907,	'ое': 0.002225369208982399,	'ня': 0.000606918875177018,	'тэ': 0.000606918875177018,
'ет': 0.00829455796075258,	'ею': 0.00040461258345134534,	'ал': 0.000606918875177018,	'эт': 0.002427675500708072,
'тя': 0.00040461258345134534,	'юж': 0.00020230629172567267,	'еу': 0.0008092251669026907,	'ощ': 0.00040461258345134534,
'ял': 0.00020230629172567267,	'эн': 0.0016184503338053813,	'уд': 0.001213837750354036,	'шу': 0.00020230629172567267,
'ад': 0.0032369006676107627,	'нь': 0.0014161440420797087,	'ер': 0.005259963584867489,	'ущ': 0.00020230629172567267,
'йа': 0.00040461258345134534,	'ью': 0.001820756625531054,	'рж': 0.00020230629172567267,	'ще': 0.001820756625531054,
'сш': 0.00020230629172567267,	'юн': 0.00040461258345134534,	'ва': 0.0074853327938498885,	'ий': 0.0020230629172567267,
'ше': 0.00020230629172567267,	'ожд': 0.001213837750354036,	'ар': 0.003439206959336435,	'йв': 0.0008092251669026907,
'тн': 0.002427675500708072,	'юю': 0.001213837750354036,	'рв': 0.000606918875177018,	'ач': 0.0016184503338053813,
'дц': 0.000606918875177018,	'йе': 0.00020230629172567267,	'ми': 0.0004046125834513453,	'чу': 0.00020230629172567267,
'ца': 0.0008092251669026907,	'ед': 0.003641513251062108,	'ьо': 0.000606918875177018,	'ме': 0.004653044709690472,
'ат': 0.004855351001416144,	'ду': 0.0028322880841594174,	'бя': 0.00040461258345134534,	'рл': 0.00020230629172567267,
'ти': 0.008901476835929597,	'уш': 0.0008092251669026907,	'ян': 0.0020230629172567267,	'тэ': 0.00020230629172567267,
'ил': 0.0038438195427877805,	'ша': 0.00020230629172567267,	'же': 0.002225369208982399,	'че': 0.005462269876593162,
'ер': 0.0020230629172567267,	'ан': 0.0014161440420797087,	'оя': 0.0008092251669026907,	'йл': 0.00040461258345134534,
'оз': 0.0020230629172567267,	'ир': 0.0032369006676107627,	'еч': 0.001820756625531054,	'дж': 0.00020230629172567267,
'эр': 0.00040461258345134534,	'зг': 0.00040461258345134534,	'кр': 0.0020230629172567267,	'ря': 0.00040461258345134534,
'ра': 0.013554521545620068,	'ьи': 0.0014161440420797087,	'уж': 0.0010115314586283633,	'яс': 0.001213837750354036,
'ао': 0.002427675500708072,	'ип': 0.0038438195427877805,	'жа': 0.000606918875177018,	'ий': 0.001820756625531054,
'ап': 0.0020230629172567267,	'га': 0.001820756625531054,	'йн': 0.00040461258345134534,	'йд': 0.0016184503338053813,
'оч': 0.0028322880841594174,	'са': 0.001213837750354036,	'сс': 0.002225369208982399,	'ия': 0.004046125834513453,
'чт': 0.001820756625531054,	'оо': 0.0008092251669026907,	'см': 0.001213837750354036,	'ео': 0.0008092251669026907,
'мн': 0.002427675500708072,	'аб': 0.0008092251669026907,	'мы': 0.001213837750354036,	'рг': 0.0008092251669026907,
'зв': 0.0032369006676107627,	'би': 0.00040461258345134534,	'ьм': 0.0026299817924337445,	'эм': 0.0028322880841594174,
'си': 0.0008092251669026907,	'ьк': 0.000606918875177018,	'ят': 0.0026299817924337445,	'тп': 0.00020230629172567267,
'им': 0.0032369006676107627,	'пт': 0.00020230629172567267,	'мк': 0.0008092251669026907,	'ри': 0.008901476835929597,
'мв': 0.004855351001416144,	'вк': 0.00020230629172567267,	'юж': 0.00020230629172567267,	'ьз': 0.00020230629172567267,
'ав': 0.004450738417964798,	'кл': 0.0008092251669026907,	'л': 0.001213837750354036,	'ий': 0.00040461258345134534,
'аж': 0.002427675500708072,	'тк': 0.0014161440420797087,	'лю': 0.001213837750354036,	'пк': 0.00040461258345134534,
'жи': 0.0016184503338053813,	'ке': 0.0008092251669026907,	'юб': 0.000606918875177018,	'лг': 0.00020230629172567267,
'об': 0.004855351001416144,	'еа': 0.00020230629172567267,	'бв': 0.00020230629172567267,	'гс': 0.00020230629172567267,
'бс': 0.00040461258345134534,	'ки': 0.004450738417964798,	'иа': 0.0010115314586283633,	'сп': 0.003439206959336435,
'ве': 0.004653044709690472,	'кт': 0.0026299817924337445,	'ум': 0.0010115314586283633,	'ся': 0.004855351001416144,
'св': 0.001820756625531054,	'то': 0.011733764920089014,	'йо': 0.000606918875177018,	'пп': 0.0010115314586283633,
'ое': 0.002225369208982399,	'ня': 0.000606918875177018,	'та': 0.000606918875177018,	'ие': 0.0004046125834513453,

'ие': 0.004046125834513453,  
'иб': 0.00020230629172567267,  
'сч': 0.000606918875177018,  
'ец': 0.001213837750354036,  
'цн': 0.00020230629172567267,  
'це': 0.0014161440420797087,  
'ын': 0.00040461258345134534,  
'ае': 0.002427675500708072,  
'бл': 0.000606918875177018,  
'лк': 0.00040461258345134534,  
'хи': 0.001213837750354036,  
'ип': 0.0008092251669026907,  
'вг': 0.00020230629172567267,  
'ву': 0.0008092251669026907,  
'уч': 0.0008092251669026907,  
'тч': 0.00040461258345134534,  
'вц': 0.00020230629172567267,  
'ау': 0.0010115314586283633,  
'ха': 0.0008092251669026907,  
'ея': 0.0008092251669026907,  
'не': 0.002225369208982399,  
'чн': 0.001213837750354036,  
'оп': 0.002225369208982399,  
'ыз': 0.00040461258345134534,  
'пн': 0.00040461258345134534,  
'зы': 0.0008092251669026907,  
'мя': 0.00040461258345134534,  
'па': 0.000606918875177018,  
'йт': 0.000606918875177018,  
'ьв': 0.001213837750354036,  
'ыр': 0.0014161440420797087,  
'жу': 0.00040461258345134534,  
'гр': 0.002427675500708072,  
'ох': 0.001213837750354036,  
'пу': 0.00040461258345134534,  
'ус': 0.001820756625531054,

'ус': 0.001820756625531054,  
'бу': 0.00020230629172567267,  
'ва': 0.0008092251669026907,  
'ры': 0.002427675500708072,  
'ья': 0.00020230629172567267,  
'ыэ': 0.00020230629172567267,  
'бх': 0.00040461258345134534,  
'ыд': 0.0008092251669026907,  
'ьт': 0.000606918875177018,  
'сх': 0.00020230629172567267,  
'эв': 0.0028322880841594174,  
'лн': 0.00040461258345134534,  
'ча': 0.00303459437588509,  
'вп': 0.000606918875177018,  
'ьп': 0.00040461258345134534,  
'пл': 0.0010115314586283633,  
'иу': 0.00020230629172567267,  
'ля': 0.0026299817924337445,  
'яи': 0.001213837750354036,  
'вд': 0.00040461258345134534,  
'лс': 0.00020230629172567267,  
'шр': 0.00020230629172567267,  
'эе': 0.00040461258345134534,  
'ыш': 0.00040461258345134534,  
'ык': 0.00040461258345134534,  
'кн': 0.001213837750354036,  
'тх': 0.00020230629172567267,  
'рс': 0.0008092251669026907,  
'мр': 0.00020230629172567267,  
'мп': 0.0014161440420797087,  
'бн': 0.00020230629172567267,  
'аш': 0.00020230629172567267,  
'щи': 0.0010115314586283633,  
'нф': 0.001213837750354036,  
'фо': 0.0014161440420797087,  
'рм': 0.0010115314586283633,

'рм': 0.0010115314586283633,  
'ац': 0.002427675500708072,  
'ци': 0.004248432126239126,  
'вя': 0.0008092251669026907,  
'яэ': 0.0008092251669026907,  
'эи': 0.0008092251669026907,  
'чи': 0.0010115314586283633,  
'ги': 0.0014161440420797087,  
'ич': 0.003641513251062108,  
'ях': 0.000606918875177018,  
'хп': 0.0008092251669026907,  
'ию': 0.0014161440420797087,  
'хр': 0.000606918875177018,  
'ул': 0.0008092251669026907,  
'вр': 0.0014161440420797087,  
'бу': 0.0010115314586283633,  
'ьм': 0.0008092251669026907,  
'ты': 0.0014161440420797087,  
'хк': 0.00020230629172567267,  
'тс': 0.0004046125834513453,  
'як': 0.00040461258345134534,  
'нк': 0.0010115314586283633,  
'оц': 0.0010115314586283633,  
'хс': 0.0008092251669026907,  
'гл': 0.00040461258345134534,  
'ьг': 0.00020230629172567267,  
'ба': 0.00040461258345134534,  
'йф': 0.00020230629172567267,  
'фл': 0.00020230629172567267,  
'рт': 0.0008092251669026907,  
'чс': 0.00020230629172567267,  
'вч': 0.00040461258345134534,  
'лт': 0.00020230629172567267,  
'му': 0.0016184503338053813,  
'уо': 0.000606918875177018,  
'йх': 0.00040461258345134534,

'йх': 0.00040461258345134534,  
'вв': 0.000606918875177018,  
'вм': 0.000606918875177018,  
'нм': 0.00020230629172567267,  
'юо': 0.0008092251669026907,  
'нэ': 0.00020230629172567267,  
'дэ': 0.00020230629172567267,  
'иг': 0.0010115314586283633,  
'уп': 0.0010115314586283633,  
'пп': 0.00020230629172567267,  
'пы': 0.00020230629172567267,  
'йу': 0.00020230629172567267,  
'уэ': 0.00020230629172567267,  
'хч': 0.00020230629172567267,  
'чл': 0.00020230629172567267,  
'бм': 0.00020230629172567267,  
'ям': 0.001213837750354036,  
'цы': 0.00040461258345134534,  
'ьч': 0.00020230629172567267,  
'оу': 0.00040461258345134534,  
'ук': 0.0008092251669026907,  
'нд': 0.000606918875177018,  
'юю': 0.0014161440420797087,  
'юш': 0.0008092251669026907,  
'мб': 0.00020230629172567267,  
'ту': 0.0059576572931418165,  
'еф': 0.00020230629172567267,  
'ьб': 0.00020230629172567267,  
'сд': 0.00040461258345134534,  
'хд': 0.0008092251669026907,  
'ик': 0.00040461258345134534,  
'пъ': 0.000606918875177018,  
'ют': 0.00020230629172567267,  
'рн': 0.0008092251669026907,  
'эб': 0.00020230629172567267,  
'кг': 0.00020230629172567267,

'зб': 0.00020230629172567267,  
'кг': 0.00020230629172567267,  
'хн': 0.00020230629172567267,  
'шн': 0.00020230629172567267,  
'мм': 0.00040461258345134534,  
'дл': 0.001213837750354036,  
'нд': 0.001213837750354036,  
'хт': 0.00020230629172567267,  
'нс': 0.000606918875177018,  
'рп': 0.00020230629172567267,  
'нц': 0.00040461258345134534,  
'ьл': 0.00020230629172567267,  
'иш': 0.00040461258345134534,  
'шь': 0.00040461258345134534,  
'ьд': 0.00040461258345134534,  
'дв': 0.00040461258345134534,  
'ьт': 0.0008092251669026907,  
'дп': 0.00020230629172567267,  
'лу': 0.000606918875177018,  
'ай': 0.00040461258345134534,  
'иэ': 0.000606918875177018,  
'тб': 0.00020230629172567267,  
'иш': 0.00020230629172567267,  
'нт': 0.00040461258345134534,  
'кс': 0.000606918875177018,  
'ую': 0.0008092251669026907,  
'юп': 0.00040461258345134534,  
'яц': 0.00020230629172567267,  
'юд': 0.00040461258345134534,  
'лл': 0.00040461258345134534,  
'аа': 0.00020230629172567267,  
'вт': 0.000606918875177018,  
'вэ': 0.0008092251669026907,  
'оф': 0.00020230629172567267,  
'фа': 0.000606918875177018,  
'яч': 0.00020230629172567267,

'яч': 0.00020230629172567267,  
'гк': 0.00020230629172567267,  
'ур': 0.005259963584867489,  
'тд': 0.000606918875177018,  
'йг': 0.00020230629172567267,  
'ьи': 0.00020230629172567267,  
'юр': 0.00020230629172567267,  
'юи': 0.00020230629172567267,  
'яю': 0.00020230629172567267,  
'яо': 0.0010115314586283633,  
'дл': 0.00020230629172567267,  
'пи': 0.00020230629172567267,  
'шо': 0.00020230629172567267,  
'юк': 0.00040461258345134534,  
'бш': 0.00020230629172567267,  
'яэ': 0.00040461258345134534,  
'эв': 0.00020230629172567267,  
'юх': 0.00020230629172567267,  
'зэ': 0.00040461258345134534,  
'йй': 0.00040461258345134534,  
'ью': 0.00040461258345134534,  
'тт': 0.00040461258345134534,  
'йч': 0.00040461258345134534,  
'юв': 0.00020230629172567267,  
'уе': 0.00040461258345134534,  
'чш': 0.00020230629172567267,  
'юе': 0.00020230629172567267,  
'хэ': 0.00020230629172567267,  
'эк': 0.00020230629172567267,  
'хк': 0.00040461258345134534,  
'тм': 0.00020230629172567267,  
'еэ': 0.00040461258345134534,  
'сш': 0.00020230629172567267,  
'кц': 0.00020230629172567267,  
'яд': 0.00020230629172567267,  
'яр': 0.00020230629172567267,

'хк': 0.00040461258345134534,  
'тм': 0.00020230629172567267,  
'еэ': 0.00040461258345134534,  
'сы': 0.00020230629172567267,  
'кц': 0.00020230629172567267,  
'яд': 0.00020230629172567267,  
'яр': 0.00020230629172567267,  
'иф': 0.000606918875177018,  
'мт': 0.0008092251669026907,  
'мэ': 0.0008092251669026907,  
'су': 0.00040461258345134534,  
'ою': 0.00020230629172567267,  
'юс': 0.00020230629172567267,  
'фи': 0.00040461258345134534,  
'ку': 0.00040461258345134534,  
'уц': 0.00020230629172567267,  
'оа': 0.00020230629172567267,  
'нг': 0.0014161440420797087,  
'хф': 0.00020230629172567267,  
'фр': 0.00020230629172567267,  
'яа': 0.00020230629172567267,  
'эс': 0.00020230629172567267,  
'аф': 0.00020230629172567267,  
'оэ': 0.00020230629172567267,  
'юч': 0.00020230629172567267,  
'тл': 0.00020230629172567267,  
'нэ': 0.00020230629172567267,  
'йя': 0.00020230629172567267,  
'яе': 0.00040461258345134534,  
'яу': 0.00020230629172567267



## CoolPinkProgram:

[illegible][illegible]



**Оцінки надлишковості:**

Значення ентропії на алфавіті з пробілом  $H_0 = 5$ ;

Значення ентропії на алфавіті без пробілу  $H_0 = 4,954$ ;

$$R = 1 - \frac{H_n}{H_0}$$

Для  $H_1$ :

$R = 0,11$  на алфавіті з пробілами;

$R = 0,108$  на алфавіті без пробілів;

Для  $H_2$ :

$R = 0,205$  на алфавіті з пробілами з перетинами;

$R = 0,213$  на алфавіті з пробілами без перетинів;

$R = 0,188$  на алфавіті без пробілів з перетинами;

$R = 0,198$  на алфавіті без пробілів без перетинів;

Для  $H^{(10)}$ :

$$0,431 \leq R \leq 0,566$$

Для  $H^{(20)}$ :

$$0,601 \leq R \leq 0,737$$

Для  $H^{(30)}$ :

$$0,667 \leq R \leq 0,769$$

**Оцінка надлишковості російської мови:**

$$R = 1 - \frac{H_\infty}{H_0} = 1 - \frac{4,447975}{5} = 0,1104$$



## Висновки

Під час виконання даної лабораторної роботи було обчислено частоти букв та біграм в тексті, а також значення ентропії за безпосереднім означенням. Як і очікувалось, отримані значення ентропій біграм на алфавіті з пробілами та без (а також з перетином та без) відрізнялись мінімально.

Обчислений результат надлишковості російської мови становить більше 11%